DISCUSSION NOTE

Unifying everything: Some remarks on simpler syntax, construction grammar, minimalism, and HPSG

STEFAN MÜLLER

Freie Universität Berlin

This discussion note compares several current linguistic theories: at the extreme ends of the current views are MINIMALIST theories, which restrict themselves to the binary branching operations Move and Merge, and SIMPLER SYNTAX, which assumes flat structures and a surface-oriented mapping between syntactic structures and grammatical functions.

I show that purely surface-oriented theories have problems in accounting for the relatedness of syntactic and morphological structures and for the iteration of valence-changing processes, and I argue for a lexical analysis, as can be found in current minimalist theories, HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR (HPSG), and CATEGORIAL GRAMMAR. I furthermore show that the Chomskyan view on label computation is problematic for several reasons and should be given up in favor of explicit accounts like the one used in HPSG. I discuss problems for the analysis of complements and specifiers in minimalist theories with special focus on Stabler's minimalist grammars. I argue that once all problems are fixed, the resulting combinatorial rules are rather similar to what HPSG does.

As various proponents of more surface-oriented theories like CONSTRUCTION GRAMMAR, SIMPLER SYNTAX, and HPSG have pointed out, two types of binary branching, headed rules are not sufficient to account for the entirety of language, which leads to the conclusion that both research directions are right to a certain extent: there is need for (constraint-based versions of) Move and Merge and there is need for special phrasal constructions.*

Keywords: passive, labeling, simpler syntax, minimalism, minimalist grammar, categorial grammar, HPSG, lexicalism

In this discussion note I want to compare several current linguistic theories: at the extreme ends of the current views are MINIMALIST theories, which restrict themselves to the binary branching operations Move and Merge (Chomsky 2008, 2013), and SIMPLER SYNTAX (Culicover & Jackendoff 2005), which assumes flat structures and a surface-oriented mapping between syntactic structures and grammatical functions. Most current linguists also understand CONSTRUCTION GRAMMAR as a theory that combines phrase structure grammar patterns with a certain meaning. Analyses that are inspired by this view can be found in frameworks like LEXICAL-FUNCTIONAL GRAMMAR (LFG; Asudeh

et al. 2008, 2013), HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR (HPSG; Haugereid 2009), and TREE-ADJOINING GRAMMAR (TAG; Kallmeyer & Osswald 2012).

Jackendoff (2011) compares the view currently assumed in minimalist circles with the constraint-based view that underlies LFG, HPSG, and simpler syntax. He argues that unification or rather constraint satisfaction is better suited to describe our linguistic capabilities than a purely binary Merge operation that is used in derivations and applied in a bottom-up fashion. He tries to support his view with a number of examples, and I agree that most of his examples are convincing. However, I show that constructions that interact with valence-changing processes and derivational morphology should not be treated in the way that Jackendoff and other proponents of phrasal analyses suggest and hence cannot be seen as data that supports phrasal approaches. I suggest here a lexical proposal instead. As a background for the lexical proposal, I assume the framework of HPSG, but of course an HPSG analysis that assumes binary branching structures is very similar to a certain kind of minimalist approach. A large part of this discussion note is devoted to current proposals by Chomsky, and I point out various problems and discuss possible solutions. As I show below, there are solutions to these problems that are free of the conceptual difficulties that haunt minimalist alternatives. The discussion of minimalism consists of a section on Merge and labeling as it was defined in Chomsky 2008, 2013, in which I show that the labeling mechanism cannot account for free relative clauses, contrary to Chomsky's claims. Furthermore, I discuss problems that result from assumptions about specifiers and complements in relation to Merge and conclude that a theory that fixes these problems is equivalent to a head-driven phrase structure grammar containing the HEAD-FILLER SCHEMA, the HEAD-SPECIFIER SCHEMA, and the HEAD-COMPLEMENT SCHEMA (only). Since the discussion of Jackendoff's proposal suggests that some of his constructions should be treated lexically, this means that there has to be a component in grammar that combines lexical items: Move and Merge in minimalist theories and the respective HPSG schemata. Since Jackendoff and others have shown that we need phrasal constructions in other parts of the grammar, Move and Merge alone are not sufficient and, hence, minimalist theories have to be extended.

So, the conclusion is that both research directions are right to a certain extent: there is need for (constraint-based versions of) Move and Merge, and there is need for special phrasal constructions.

1. STORED PHRASAL CONSTRUCTIONS AND INTERACTION WITH MORPHOLOGY, VALENCE, AND COORDINATION. Jackendoff (2011:610) argues for a view of grammar that allows for complex linguistic objects to be stored in the lexicon. The background for this is that certain phrasal configurations are associated with a meaning that does not come from the individual parts, but is associated with the phrase as a whole (Goldberg 1995, 2006, Goldberg & Jackendoff 2004, Jackendoff 2008, Jacobs 2008). Among other examples, Jackendoff gives the two in 1.[1]

    (1) a. The bus rumbled around the corner.
               $[_{VP} \text{PP}]$ = 'go PP in such a way to make a V-ing sound'
      b. student after student
            $[_{NP/advP} \text{N-P-N}]$

Jackendoff (2008) argues, correctly in my opinion, that an analysis that treats examples like 1b as instances of headed structures has to make technical assumptions that are not motivated by the data. The analysis that was suggested by G. Müller (2011:§3.3) suffers

---

[1] N-P-N constructions were first discussed by Williams (1994:18–20).

from the very shortcomings that were identified by Jackendoff. So, a phrasal construction is needed for 1b. But 1a is different. I think it can be shown that the basic analysis of all phenomena interacting with valence-alternation processes and/or derivational morphology should be lexical in nature, rather than connected to a phrasal skeleton. I made this point earlier in Müller 2006 with respect to resultative constructions (see also Jacobs 2009) and want to repeat the crucial points here: the important generalization about the *rumble* cases involves the dependency between a lexical head and a PP, not the VP configuration they appear in. To see this, consider the German analog to this construction: as the example in 2b shows, it is possible to have the verbal stem *quietsch-* 'squeak' inside a combinatorial nominal derivation.

(2) a. dass die Straßenbahnen um   die Ecke quietschen
        that the trams   around the corner squeak
          'that the trams squeak around the corner'

    b. wegen   des Um-die-Ecke-Gequietsches   der   Straßenbahnen
        because.of the around-the-corner-squeaking.GEN of.the trams
          'because of the (annoying) squeaking around the corner of the trams'

This *Ge- -e*-derivation is a standard derivation that is independent of the *rumble* construction. If the meaning of 2b is to be tied to a phrasal configuration, that structure would be 3.

(3) [$_N$ PP [$_N$ [$_{N\text{-stem}}$ ge- V-stem -e] -s]]

So, we have a deeply embedded verbal stem that undergoes *Ge- -e*-derivation to form a nominal stem. This stem has to be inflected to form a noun, and, in a further step, the combination of the noun with the PP takes place. The semantics of the whole construction is embedded under the semantics of the *Ge- -e* circumfix. *Ge- -e* comes with a connotation that the V-ing is habitual and unpleasant (Lüdeling 2001:§3.4.3). The scope relations are most straightforwardly captured in a lexical analysis:[2] a lexical rule applies to a verb and licenses an item that selects for the PP and contributes the appropriate meaning.[3] This lexical item can enter the *Ge- -e*-derivation, and the resulting nominal can be combined with a prenominal PP and the postnominal genitive. See Müller 2002:§6.2.5 and Müller 2003c for a fully worked-out analysis of derivational morphology along the lines suggested above.

The *Ge- -e*-derivation + prenominal constituent is not restricted to the *rumble* verb construction, but follows a normal pattern of German syntax/morphology. Further examples from the DECOW2012 corpus are shown in 4.

---

[2] Cases like 4c and 4d involve adjunct PPs. The scope relations could be captured by an analysis along the lines of van Noord & Bouma 1994 or Bouma et al. 2001, which suggest the lexical introduction of adjuncts. An alternative would be Egg's analysis of Dowty's *open again* examples that accounts for sublexical scopings in an underspecification approach (Dowty 1979:§5.6, Egg 1999). So, while the lexical analysis does account for the scope facts directly, this is not the only argument for it: the fact that it is difficult to see how 1a and 3 can be related is not affected by the possibility of sublexical scoping.

[3] There are several views on lexical rules (Jackendoff 1975). I assume a version of lexical rules that corresponds to a unary branching structure. For a formalization of this view on lexical rules see Briscoe & Copestake 1999 and Meurers 2001. Briscoe and Copestake show how the various variants of the ditransitive construction can be treated lexically. See Müller 2013c:§11.10 for the relation between lexical rules, empty elements, and transformations.

Goldberg (2013) discusses lexical and phrasal approaches. She differentiates the lexical approaches into ones that assume lexical rules and ones with lexical templates. She argues against lexical rule-based approaches, but most of her points are compatible with what she calls lexical templates. The view on lexical rules that is standard in HPSG nowadays corresponds to Goldberg's lexical templates.

(4) a. an-den-Hals-Gewerfe
    toward-the-neck-throwing
      'the repeated annoying throwing oneself around the neck of somebody
        else'
  b. aus-dem-Bett-Gequäle
    out-of-the-bed-forcing
      'the repeated unpleasant forcing oneself out of bed'
  c. im-Kreis-Gelaufe
    in.the-circle-running
      'the repeated annoying running in circles'
  d. auf-dem-Tisch-Getrommle
    on-the-table-drumming
      'the repeated annoying drumming on the table'

Because of the possibility of realizing these verb-PP combinations both in a VP config-
uration and in the *Ge- -e*-derivation, phrasal approaches that refer to syntactic structure
have to mention two unrelated structures, and it is unclear how the commonalities be-
tween the structures—that is, the dependency between the involved items—can be cap-
tured. GENERALIZED PHRASE STRUCTURE GRAMMAR-like METARULES (GPSG; Gazdar et
al. 1985) or transformations seem to be the only way out. Note, though, that GPSG
metarules operate on phrase structure rules, that is, local trees. In contrast, 3 involves a
rather complex structure.

  I am fully aware of the fact that languages differ and that authors who do not believe
in universal grammar (UG) are not necessarily concerned with capturing crosslinguistic
commonalities. After all, children have to acquire their language from the input in their
environment (Tomasello 2003, Freudenthal et al. 2006, 2007, Bod 2009), and the exis-
tence of English speakers on an island far away does not influence the acquisition of
German. If we can choose between several theoretical approaches, however, we should
take the one that can capture crosslinguistic generalizations. This is the case for the lex-
ical approach. It captures the commonalities between German and English with respect
to 1a and 2a without saying anything about the particular realization of the dependents.

  Another problem of phrasal approaches was mentioned in Müller 2006:§5.2: if one
assumes lexical integrity (Bresnan & Mchombo 1995), it cannot be explained why the
formation of adjectival participles from monovalent verbs is possible when they are
part of a resultative construction, while it is excluded without a result predicate. Exam-
ples 5 and 6 show the relevant examples: the examples in 5 show that the adjectival par-
ticiple can be formed only if the verb is transitive. Verbs like *tanzen* 'to dance' that do
not take an object do not form adjectival participles.[4] As 6 shows, however, adjectival
participles can appear in resultative constructions.

(5) a. Der Mann liebt die Frau.
      the man loves the woman
  b. die geliebte Frau
      the loved   woman
        'the beloved woman'
  c. *Der Mann tanzt   die Schuhe.
      the man   dances the shoes

---

  [4] *tanzen* 'to dance' can take a cognate object like *the waltz*. I am talking about the monovalent version of
tanzen here.

    d. *die getanzten Schuhe
      the danced   shoes
(6) a. Der Mann tanzt die Schuhe blutig / in   Stücke.
     the man dances the shoes bloody into pieces
   b. die blutig / in   Stücke getanzten Schuhe
     the bloody into pieces danced   shoes

This means that the morphology has to have access to the fact that there will be an accusative object, that is, that *tanz-* is used in a resultative construction. As a consequence of this, the accusative object cannot be licensed by a phrasal construction into which a fully inflected item enters, but it has to be represented at the lexical level. This is a variant of a very old argument for the lexical treatment of passive by Dowty (1978:412) and Bresnan (1982:21). If the resultative construction is treated lexically, a special lexical item exists that selects for a subject, accusative object, and result predicate. This lexical item can undergo passive participle inflection, adjective derivation, and adjective inflection and can be used in 6b, or it can be inflected and used in 6a.

Before we turn to active/passive alternations, I want to discuss another major problem for accounts that assume that argument structure can be extended by inserting heads into a phrasal configuration: it has gone unnoticed until now that such proposals have difficulties with symmetric coordination. A generalization about coordination is that two constituents that have compatible syntactic properties can be coordinated and that the result of the coordination is an object that has the syntactic properties of the linguistic objects that are coordinated. Some examples are shown in 7: in 7a we have a case of the coordination of two lexical verbs. The coordination *know and like* behaves like the coordinated simplex verbs: it takes a subject and an object. Similarly, two sentences with a missing object are coordinated in 7b, and the result is a sentence with a missing object.

(7) a. I know and like this record.
   b. Bagels, I like and Ellison hates.   (Pollard & Sag 1994:205)

The German examples in 8 show that the case requirements of the involved verbs must be observed. In 8b,c the coordinated verbs require accusative and dative, respectively, and since the case requirements are incompatible with unambiguously case-marked nouns, both of these examples are out.

(8) a. Ich kenne und unterstütze diesen Mann.
     I   know and support   this   man.ACC
   b. *Ich kenne und helfe diesen Mann.
      I   know and help this   man.ACC
   c. *Ich kenne und helfe diesem Mann.
      I   know and help this   man.DAT

It is interesting to see that the coordination of simplex ditransitive verbs with verbs that have additional arguments that are licensed by an argument structure construction is possible. As an example, let us have a look at the verb *bake*. The verb *bake* is usually a strictly transitive verb, but a benefactive argument can be realized as in 9.

(9) Sally baked her sister a cake.

Goldberg (1995:§6.2) argues that the benefactive argument in 9 is provided by a phrasal construction rather than lexically licensed.

Some examples for the coordination of *bake* taking a benefactive argument and the simplex ditransitive verb *give* are provided in 10.

(10) a. My sisters just baked and gave me a nutella cupcake with mint chocolate chip ice-cream in the middle and milk chocolate frosting on top. – :)[5]
   b. Randy baked and gave me a loaf of bread!!![6]

These sentences could be instances of RIGHT NODE RAISING (Bresnan 1974, Abbott 1976). In such a situation *bake* would have been used without a benefactive object, as in 11.

(11) baked a cake and gave me a cake

Furthermore, as Mary Dalrymple pointed out to me, the examples in 10 could have a structure with an intransitive *bake*. However, such readings are not available for the German counterpart in 12.

(12) ich hab ihr jetzt diese Ladung Muffins mit den Herzchen drauf
    I   have her now this   load   muffins with the little.heart here.on
        gebacken und gegeben.[7]
        baked   and given
    'I now baked and gave her this load of muffins with the little heart on top.'

The problem with such sentences now is the following: if it is the verbs that are coordinated and if the coordinated items have to have compatible syntactic properties like valence, this means that *gebacken* and *gegeben* have to have the same valence properties. This is accounted for in a lexical approach since in the lexical approach the transitive version of the creation verb *bake* licenses a ditransitive verb that can be coordinated with *give*. In the phrasal approach, however, the verb *bake* has two argument roles and is not compatible with the verb *give*, which has three argument roles. In the phrasal model, *bake* can only realize three arguments when it enters the respective argument structure construction, but in sentences like 12 it is not *bake* alone that enters the phrasal syntax, but rather the combination of *baked* and *given*.[8] But the verbs are incompatible as far as the semantic roles are concerned. To fix this, one would have to assume that the semantic roles that are required from the phrase *baked and given* percolate down to the verbs and that they are therefore compatible. But this would amount to saying that there are several verb senses for *baked*, something that proponents of phrasal analyses try to avoid (Goldberg 1995:§1.4.1, 2013:§3.1.1).

Before turning to minimalist theories I want to look at another group of phenomena that should not be treated on the phrasal level: valence-changing processes. Culicover and Jackendoff (2005:Ch. 6.3) suggest that passive is an alternative mapping of grammatical functions onto a phrase structure configuration. Their proposal is just an instance of a set of proposals that assume that there is a certain core representation of verbs and that this core representation can be unified with either a passive or an active description (see for instance Candito 1996, Kay & Fillmore 1999:12, Koenig 1999:Ch. 3, Davis & Koenig 2000, Kordoni 2001, Michaelis & Ruppenhofer 2001:Ch. 4, Clément & Kinyon 2003:188 for proposals in construction grammar, HPSG, and TAG). Culicover and Jackendoff discuss the example in Figure 1.

---

[5] http://bambambambii.tumblr.com/post/809470379, accessed June 5, 2012.

[6] http://marilynmillersstudio.blogspot.com/2012/03/floral-celebration-of-friendships.html, accessed July 16, 2012.

[7] http://www.musiker-board.de/diverses-ot/35977-die-liebe-637-print.html, accessed June 8, 2012.

[8] See also Wechsler 2008:505 for a similar coordination argument regarding accounts like the one by Marantz (1997) that assume decomposition in syntax.

DESIRE(BILL$_2$,   [SANDWICH; DEF]$_3$)
                | |
               GF$_2$ GF$_3$
                | |
[$_S$ NP$_2$ [$_{VP}$ V$_1$ NP$_3$]]

          Bill   desires   the sandwich.

DESIRE(BILL$_2$,   [SANDWICH; DEF]$_3$)
                | |
               GF$_2$ GF$_3$

[$_S$ NP$_3$ [$_{VP}$ V$_1$ by NP$_2$]]

          the sandwich   is desired   by Bill.
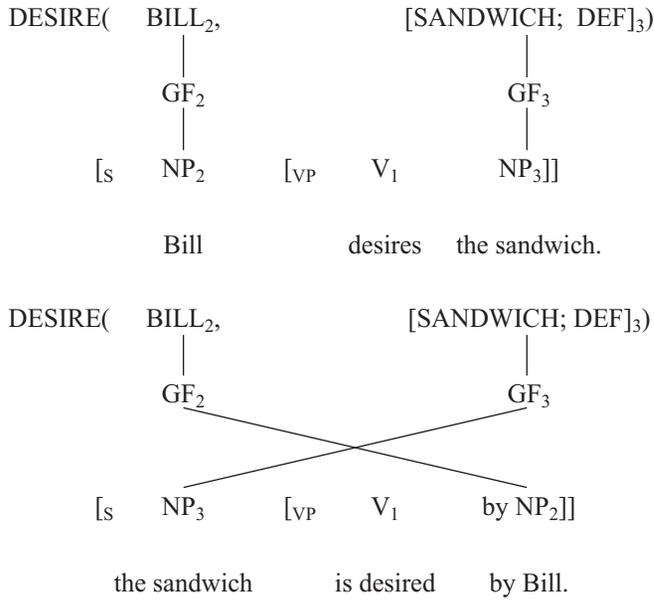
FIGURE 1. Active and passive mappings according to Culicover and Jackendoff (simplified).

The mapping in the upper part of the figure comes about by unification with the constraints on active sentences, and the mapping in the lower part of the structure is the result of unifying in the constraints for passive sentences. As was pointed out in Müller 2006:§4, 2007a:§5, languages that allow multiple application of valence-changing processes pose a problem for such approaches. Take for instance passive: languages like Lithuanian (Timberlake 1982:§5), Irish (Noonan 1994), and Turkish (Özkaragöz 1986) allow for impersonalizations of personal passives. The cited authors discussed their data as instances of double passivization, but it was argued by Blevins (2003) that these and similar examples from other languages are impersonal constructions that can be combined with personal passives. Özkaragöz's Turkish examples are used in 13 for illustration (1986:77).

(13) a. Bu şato-da   bo-ul-un-ur.
             this chateau-LOC strangle-PASS-PASS-AOR
                'One is strangled (by one) in this chateau.'
      b. Bu oda-da   döv-ül-ün-ür.
             this room-LOC hit-PASS-PASS-AOR
                'One is beaten (by one) in this room.'
      c. Harp-te   vur-ul-un-ur.
             war-LOC shoot-PASS-PASS-AOR
                'One is shot (by one) in war.'

-*In*, -*n*, and -*Il* are allomorphs of the passive morpheme. According to Özkaragöz, the data is best captured by an analysis that assumes that the passive applies to a passivized transitive verb and hence results in an impersonal passive. The examples in 13 use his glossing. If Blevins's analysis is correct, the last morpheme glossed as passive in 13 has to be glossed as impersonal. Approaches that assume that the personal passive is the unification of some general structure with some passive-specific structure will not be able to capture double passivization or passivization + impersonalization, since they

committed to a certain structure too early. The problem for nontransformational approaches that state syntactic structure for the passive is that such a structure, once stated, cannot be modified. That is, we said that the underlying object is the subject in the passive sentence. But in order to get the double passivization/passivization + impersonalization, we have to suppress this argument as well. What is needed is some sort of process (or description) that takes a representation and relates it to another representation with a suppressed subject. This representation is related to a third representation that again suppresses the subject, resulting in an impersonal sentence. In order to do this, one needs different strata as in relational grammar (Timberlake 1982, Özkaragöz 1986), metarules (Gazdar et al. 1985), lexical rules (Dowty 1978:412, 2003:§3.4, Bresnan 1982, Pollard & Sag 1987, Blevins 2003, Müller 2003b), transformations (Chomsky 1981), or just a morpheme-based morphological analysis that results in items with different valence properties when the passivization morpheme is combined with a head. For constructional approaches, which exclude transformations of complex tree objects, the lexical treatment seems to be the only option.

This section suggested an analysis that relies on a rich lexicon and uses syntactic schemata to combine lexical items with their dependents. In Müller 2013a,e, I suggest using binary branching rules for head-argument and head-filler combinations.[9] The following section discusses Move and Merge, the analogs in recent minimalist work.

2. Binary merge in a constraint-based setting: merge and hpsg. As was mentioned in the introduction, Jackendoff (2011) argued for constraint-based theories and against minimalist approaches that use a purely binary Merge operation in derivations in a bottom-up fashion. See also Labelle 2007 on bottom-up computational systems and Sag & Wasow 2011 on requirements for performance-compatible competence grammars. Jackendoff (2011:600) mentions that Merge can be recast in a constraint-based format. In the following I want to show that there are problems with Chomsky's definition of Merge and labeling. If these problems are fixed, I argue, one arrives at a version of HPSG with binary branching structures. Since HPSG is a theory with model-theoretic foundations (King 1999, Pollard 1999, Richter 2007), and hence a constraint-based theory, the respective HPSG schemata can perform the role of Merge in minimalism and simultaneously meet Jackendoff's desiderata.

2.1. Labeling. Chomsky (2008) distinguishes two basic operations Move and Merge, also called internal Merge and external Merge. Internal and external Merge combine two syntactic objects α and β into a bigger syntactic object that is represented as a set {α, β}. α and β can be either lexical items or internally complex syntactic objects. Internal Merge moves a part of an object to its periphery. The result of internally merging an element is a set {α, β}, where α was a part of β. External Merge also produces a set with two elements. However, two independent objects are merged. The objects that are created by Merge have a certain category (a set of features). For instance, if we combine the elements α and β, we get {l, {α, β}}, where l is the category of the resulting object. This category is also called a 'label'. Since it is assumed that all constituents are headed, the category that is assigned to {α, β} has to be either the category of α or the category of β. Chomsky (2008:145) discusses the two rules in 14 for the determination of the label of a set.

---

[9] See also Gunji 1986 for binary branching in an HPSG fragment of Japanese, Kiss 1995, and Müller 1996a on binary branching for HPSG fragments of German, and Bender et al. 2002 for binary branching in another crosslinguistic HPSG grammar project.

(14) a. In {H, α}, H an LI, H is the label.
      b. If α is internally merged to β forming {α, β}, then the label of β is the label of {α, β}.

As Chomsky notes, these rules are not unproblematic since the label is not uniquely determined in all cases. An example is the combination of two lexical elements. If both H and α in 14a are lexical items (LI), then both H and α can be the label of the resulting structure. Chomsky notices that this could result in deviant structures, but claims that this concern is unproblematic and ignores it. Chomsky delivered a treatment of the combination of two lexical items in 2013. The solution to the problem is to assume that all combinations of lexical elements consist of a functional element and a root (Marantz 1997, Borer 2005). Per definition, roots are not considered as labels,[10] and hence the category of the functional element determines the category of the combination (2013:47). Such an analysis can only be rejected: the goal of the minimalist program (MP) is to simplify the theoretical proposals to such an extent that the models of language acquisition and language evolution become plausible, but in order to simplify basic concepts it is stipulated that a noun cannot simply be a noun but needs a functional element to tell the noun what category it has. Given that the whole point of Chomsky's BARE PHRASE STRUCTURE was the elimination of the unary branching structures in X′ theory, it is unclear why they are reintroduced now through the back door, only more complex with an additional empty element.[11] Theories like CATEGORIAL GRAMMAR (CG) and HPSG can combine lexical items directly without assuming any auxiliary projections or empty elements. See also Rauh 2013 for a comparison of the treatment of syntactic categories in earlier versions of transformational grammar, HPSG, construction grammar, role and reference grammar, and root-based neo-constructivist proposals like the one assumed by Chomsky (2013). Rauh concludes that the direct connection of syntactic and semantic information is needed and that the neo-constructivism of Marantz and Borer has to be rejected. For further criticism of neo-constructivist approaches see Wechsler 2008 and Müller & Wechsler 2014:§7. The combination of a pronoun with a verbal projection poses a problem that is related to what has been said above. In the analysis of *He left*, the pronoun *he* is a lexical element and hence would be responsible for the label of *He left*, since *left* is an internally complex verbal projection in minimalist theories. The result would be a nominal label rather than a verbal one. To circumvent this problem, Chomsky (2013:46) assumes that *he* has a complex internal structure: 'perhaps D-pro', that is, *he* is (perhaps) composed out of an invisible determiner and a pronoun.

The case in which two non-LIs are externally merged (for instance, a nominal and a verbal phrase) is not discussed in Chomsky 2008. Chomsky (2013:43–44) suggests that a phrase XP is irrelevant for the labeling of {XP, YP} if XP is moved (or rather, copied, in the copy theory of movement) in a further step. Chomsky assumes that one of two phrases in a {XP, YP} combination has to move, since otherwise labeling would be impossible (2013:45).[12] The following coordination example illustrates this: Chomsky as-

---

[10] Another category that is excluded as label per definition is *Conj*, which stands for conjunction (Chomsky 2013:45–46). This is a stipulation that is needed to get coordination to work. See below.

[11] The old X′ rule in (ia) corresponds to the binary combination in (ib).

    (i) a. N′ → N
        b. N → N-func root

In (ia) a lexical noun is projected to an N′, and in (ib) a root is combined with a functional nominal head into a nominal category.

[12] His explanation is contradictory: on p. 11 Chomsky assumes that a label of a combination of two entities with the same category is this category. But in his treatment of coordination, he assumes that one of the conjuncts has to be raised, since otherwise the complete structure could not be labeled.

sumes that the expression *Z and W* is analyzed as follows. First, Z and W are merged. This expression is combined with Conj (15a), and in the next step Z is raised (15b).

(15) a. [$_\alpha$ Conj [$_\beta$ Z W]]
     b. [$_\gamma$ Z [$_\alpha$ Conj [$_\beta$ Z W]]

Since Z in β is only a copy, it does not count for labeling and β can get the label of W. It is stipulated for the combination of Z and α that Conj cannot be the label, and hence the label of the complete structure is Z.[13]

A special case that is discussed by Chomsky is the internal Merge of an LI α with a non-LI β. According to rule 14a, the label would be α. According to 14b, the label would be β (see also Donati 2006). Chomsky discusses the combination of the pronoun *what* with *you wrote* as an example.

(16) what [C [you wrote *t*]]

When the label is determined according to 14b, one gets a syntactic object that would be called a CP in the government and binding (GB) framework; since this CP is, moreover, interrogative, it can function as the complement of *wonder*, as in 17a. When the label is determined according to 14a, one gets an object that can function as the accusative object of *read* in 17b—that is, something that corresponds to a DP in GB terminology.

(17) a. I wonder what you wrote.
     b. I read what you wrote.

What you wrote in 17b is a so-called free relative clause.

Chomsky's approach to free relative clauses is interesting but is unable to describe the phenomenon in full breadth. The problem is that the phrase that contains the relative pronoun may be complex (contrary to Donati's claims; see also Citko 2008:930–32).[14] An English example from Bresnan & Grimshaw 1978:333 is provided in 18, and German examples from Bausewein 1990:155 and Müller 1999b:78 are given in 19.

(18) I'll read [whichever book] you give me.
(19) a. Ihr   könnt beginnen, [mit **wem**] ihr wollt.
        you can   start   with whom you want
          'You can start with whoever you like.'   (Bausewein 1990:155)
     b. [**Wessen** Birne] noch halbwegs in der Fassung steckt, pflegt solcherlei
         whose nut   yet halfway in the holder is   uses such
            Erloschene zu meiden;
            extinct   to avoid
          'Those who still have their wits half way about them tend to avoid such
            vacant characters;'   (Thomas Gsella, *taz*, February 12, 1997, p. 20)

---

[13] As Bob Borsley (p.c. 2013) pointed out to me, this makes wrong predictions for coordinations of two singular noun phrases with *and*, since the result of the coordination is a plural NP and not a singular one like the first conjunct. Theories like HPSG can capture this by grouping features in bundles that can be shared in coordinated structures (syntactic features and nonlocal features; see Pollard & Sag 1994:202). Furthermore, the whole account cannot explain why (ib) is ruled out.

(i) a. both Kim and Lee
    b. *both Kim or Lee

The information about the conjunction has to be part of the representation for *or Lee* in order to be able to contrast it with *and Lee*.

A further problem is that the label of α should be the label of W since *Conj* does not count for label determination. This would lead to a situation in which we have to choose between Z and W to determine the label of γ. Following Chomsky's logic, either Z or W would have to move on to make γ labelable. Chomsky mentions this problem in footnote 40, but does not provide a solution.

[14] Chomsky (2013:47) admits that there are many open questions as far as the labeling in free relative clauses is concerned and hence admits that there remain many open questions with labeling as such.

    c. [**Wessen** Schuhe] „danach" besprenkelt sind, hat keinen Baum gefunden
        whose shoes   after.that speckled   are has no   tree found
        und war nicht zu einem Bogen in der Lage.
        and was not to a   bow in the position
        'If you end up with spattered shoes afterwards it is because you
        couldn't find a tree and you were incapable of peeing in an arc.'
                                           (*taz*, August 8/9, 1998, p. xii)

Since *wessen Schuhe* 'whose shoes' is not a lexical item, rule 14b has to be applied, provided no additional rules are assumed to deal with such cases. This means that the whole free relative clause *wessen Schuhe danach besprenkelt sind* is labeled as CP. For the free relatives in 18 and 19 the labeling as a CP is an unwanted result, since they function as subjects or objects of the matrix predicates and hence should be labeled DP. Since *wessen Schuhe* is a complex phrase and not a lexical item, however, 14a does not apply and hence there is no analysis of the free relative clause as a DP. Therefore, it seems one must return to something like the GB analysis proposed by Gross and van Riemsdijk (1981), at least for the German examples. Gross and van Riemsdijk assume that free relatives consist of an empty noun that is modified by the relative clause like a normal noun. With such an approach the complexity of the relative phrase is irrelevant. Only the empty head is relevant for labeling the whole phrase.[15] But once empty heads are countenanced in the analysis, the application of 14a to 16 is undesirable since the application would result in two analyses for 17b: one with the empty nominal head and one in which 16 is labeled as NP directly. One might argue that in the case of several possible derivations, the most economical one wins, but the assumption of transderivational constraints leads to undesired consequences (Pullum 2013:§5).

---

[15] Assuming an empty head is problematic since it may be used as an argument only in those cases in which an adjunct (the relative clause) is present (Müller 1999b:97). See also Ott 2011:187 for a later rediscovery of this problem. It can be solved in HPSG by assuming a unary projection that projects the appropriate category from a relative clause. I also use the unary projection to analyze so-called nonmatching free relative clauses. In constructions with nonmatching free relative clauses, the relative clause fills an argument slot that does not correspond to the properties of the relative phrase (Bausewein 1990). Bausewein discusses the following example, in which the relative phrase is a PP but the free relative fills the accusative slot of *kocht* 'cooks'.

   (i) Sie kocht, worauf   sie Appetit hat.
        she cooks where.on she appetite has
          'She cooks what she feels like eating.'

See Müller 1999b:60–62 for corpus examples. Minimalist theories do not employ unary projections. Ott (2011) develops an analysis in which the category of the relative phrase is projected, but he does not have a solution for nonmatching free relative clauses (p. 187). The same is true for Citko's analysis, in which an internally merged XP can provide the label.

   Many other proposals for labeling, or rather nonlabeling, exist. For instance, some minimalists want to eliminate labeling altogether and argue for a label-free syntax. As was pointed out by Osborne, Putnam, and Groß (2011), such analyses bring minimalism closer to DEPENDENCY GRAMMAR. It is unclear how any of these models could deal with nonmatching free relative clauses. Groß and Osborne (2009:§5.3.3) provide an analysis of free relatives in their version of dependency grammar, but deny the existence of nonmatching ones (p. 78). They suggest an analysis in which the relative phrase is the root/label of the free relative clause, and hence they have the same problem as minimalist proposals have with nonmatching free relative clauses. As Groß and Osborne (2009:73) and Osborne and colleagues (2011:327) state: empty heads are usually not assumed in (their version of) dependency grammar. Neither are unary branching projections. This seems to make it impossible to state that free relative clauses with a relative phrase YP can function as XP provided XP is a category that is higher in the obliqueness hierarchy of Keenan and Comrie (1977), a generalization that was discovered by Bausewein (1990) (see also Müller 1999b:60–62 and Vogel 2001). In order to be able to express the relevant facts, an element or a label has to exist that is different from the label of *worauf* in (i).

Chomsky (2013) abandons the labeling condition in 14b and replaces it with general labeling rules that hold for both internal and external Merge of two phrases. He distinguishes two cases. In the first case, labeling becomes possible since one of the two phrases from the set {XP, YP} is moved away. We have already discussed this case. Chomsky writes about the other case: 'X and Y are identical in a relevant respect, providing the same label, which can be taken as the label of the SO' (2013:43). He sketches an analysis of interrogative clauses on p. 45, in which the interrogative phrase has a Q feature and the remaining sentence from which the Q phrase was extracted has a Q feature as well. Since the two constituents share this property, the label of the complete clause will be Q. This kind of labeling will 'perhaps' also be used for labeling normal sentences consisting of a subject and a verb phrase agreeing in person and number. These features would be responsible for the label of the sentence. The exact details are not worked out but almost certainly will be more complex than 14b.

A property that is inherent in both Chomsky 2005 and Chomsky 2013 is that the label is exclusively determined from one of the merged objects. As Bob Borsley pointed out to me, this is problematic for interrogative/relative phrases like 20.

    (20) with whom

The phrase in 20 is both a prepositional phrase (because the first word is a preposition) and an interrogative/relative phrase (because the second word is an interrogative/relative word. So, what is needed for the correct labeling of PPs like the one in 20 is a well-defined way of percolating different properties from daughters to the mother node.[16]

Summarizing, we can say that the labeling that was introduced to simplify the theory and reduce the amount of innate, language-specific knowledge that has to be assumed can only be made to work with a considerable number of stipulations. For instance, the combination of lexical elements requires the assumption of empty functional heads, whose only purpose is the determination of the syntactic category of a certain lexical element. If this corresponded to linguistic reality, knowledge about labeling, the respective functional categories, and information about those categories that have to be ignored for the labeling would have to be part of innate, language-specific knowledge, and nothing would be gained. We would be left with bizarre analyses with an enormous degree of complexity without having made progress in the minimalist direction. Furthermore, we have empirical problems and a large number of unsolved cases. The conclusion is that the label of a binary combination should not be determined in the ways suggested by Chomsky (2008, 2013). An alternative way to compute the label is to use the functor of a functor argument structure as the label (Berwick & Epstein 1995:145).

---

[16] HPSG solves this problem by distinguishing head features including part-of-speech information and nonlocal features containing information about extraction and interrogative/relative elements. Head features are projected from the head, and the nonlocal features of a mother node are the union of the nonlocal features of the daughters minus those that are bound off by certain heads or in certain configurations.

Citko (2008:926) suggests an analysis in which both daughters can contribute to the mother node. The result is a complex label like {P, {D, N}}. This is a highly complex data structure, and Citko does not provide any information on how the relevant information that it contains is accessed. Is an object with the label {P, {D, N}} a P, a D, or an N? One could say that P has priority since it is in the least embedded set, but D and N are in one set. What about conflicting features? How does a preposition that selects for a DP decide whether {D, N} is a D or an N? In any case, it is clear that a formalization will involve recursive relations that dig out elements of subsets in order to access their features. This adds to the overall complexity of the proposal and is clearly dispreferred over the HPSG solution, which provides one feature structure per linguistic object.

This is the approach taken by CG (Ajdukiewicz 1935, Steedman 2000) and in Stabler's minimalist grammars (2011).[17] Stabler's formalization of Merge is discussed in §2.3.

2.2. SPECIFIERS, COMPLEMENTS, AND THE REMAINS OF X′ THEORY. Chomsky (2008: 146) assumes that every head has exactly one complement but an arbitrary number of specifiers. In standard X′ theory, the restriction that there can be at most one complement followed from the general X′ schema and the assumption that structures are at most binary branching: in standard X′ theory, a lexical head was combined with all of its complements to form an X′. If there are at most two daughters in a phrase, it follows that there can be only one complement (sentences with ditransitive verbs have been analyzed with an empty head licensing an additional argument; see Larson 1988, and Müller & Wechsler 2014:§7 for a critical assesment of approaches involving little v). In standard X′ theory there was just one specifier. This restriction is now abandoned. Chomsky writes that the distinction between specifier and complement can now be derived from the order in which elements are merged with their head: elements that are FIRST-MERGED are complements, and all others—those that are LATER-MERGED—are specifiers.

Such an approach is problematic for sentences with monovalent verbs: according to Chomsky's proposal, subjects of monovalent verbs would not be specifiers but complements.[18] This problem is discussed in more detail in §2.3.

Apart from this, theories assuming that syntactic objects merged with word groups are specifiers do not allow for analyses in which two lexical verbs are directly coordinated, as in 21.[19]

    (21) He [knows and loves] this record.

For example, in an analysis suggested by Steedman (1991:264), *and* (being the head) is first merged with *loves* and then *knows*. The result of this combination is a complex ob-

---

[17] For the CG approach to work, it is necessary to assign the category x/x to an adjunct, where x stands for the category of the head to which the adjunct attaches. For instance, an adjective combines with a nominal object to form a nominal object. Therefore its category is n/n rather than adj. Similarly, Stabler's approach does not extend to adjuncts unless he is willing to assign the category noun to attributive adjectives. One way out of this problem is to assume a special combination operation for adjuncts and their heads (see Frey & Gärtner 2002:§3.2). Such a combination operation is equivalent to the head-adjunct schema of HPSG.

[18] Pauline Jacobson (p.c. 2013) pointed out that the problem with intransitive verbs could be solved by assuming that the last-merged element is the specifier and all non-last-merged elements are complements. This would solve the problems with intransitive verbs and with the coordination of verbs in 21, but it would not solve the problem of coordination in head-final languages as in 24. Furthermore, current minimalist approaches make use of multiple specifiers, and this would be incompatible with the Jacobsonian proposal unless one would be willing to state more complicated restrictions on the status of non-first-merged elements.

[19] Chomsky (2013:46) suggests the coordination analysis in 15: according to this analysis, the verbs would be merged directly, and one of the verbs would be moved around the conjunction in a later step of the derivation. As was mentioned in the previous section, such analyses do not contribute to the goal of making minimal assumptions about innate, language-specific knowledge since it is absolutely unclear how such an analysis of coordination would be acquired by language learners. Hence, I do not consider this coordination analysis here.

Another innovation of Chomsky's 2013 paper is that he eliminates the concept of specifier. He writes in footnote 27 on page 43: 'There is a large and instructive literature on problems with Specifiers, but if the reasoning here is correct, they do not exist and the problems are unformulable'. This is correct, but this also means that everything that has been explained with reference to the notion of specifier in the minimalist framework until now does not have an explanation any longer. If one follows Chomsky's suggestion, a large part of the linguistic research of the past years becomes worthless and has to be redone. Chomsky did not commit himself to a particular view on linearization in his earlier work, but somehow one has to ensure that the entities that were called specifier are realized in a position in which constituents are realized that used to be called specifier. This means that the following remarks are relevant even under current Chomskyan assumptions.

ject that has the same syntactic properties as the combined parts: the result is a complex verb that needs a subject and an object. After the combination of the conjunction with the two verbs, the result has to be combined with *this record* and *he*. *This record* behaves in all relevant respects like a complement. Following Chomsky's definition, however, it should be a specifier, since it is combined with the third application of Merge. The consequences are unclear. Chomsky assumes that Merge does not specify constituent order. According to him the linearization happens at the level of phonological form (PF). The restrictions that hold there are not described in his recent papers. However, if the categorization as complement or specifier plays a role for linearization as in Kayne's work (2011:2, 12) and in Stabler's proposal (see §2.3), *this record* would have to be serialized before *knows and loves*, contrary to the facts. This means that a CG-like analysis of coordination is not viable, and the only remaining option would seem to be assuming that *knows* is combined with an object and that two VPs are then coordinated. Kayne (1994:61, 67) follows Wexler and Culicover (1980:303) in suggesting such an analysis and assumes that the object in the first VP is deleted. Borsley (2005:471), however, showed that such an analysis makes wrong predictions, since 22a would be derived from 22b, although these sentences differ in meaning.[20]

    (22) a. Hobs whistled and hummed the same tune.
       b. Hobs whistled the same tune and hummed the same tune.

Since semantic interpretation cannot see processes like deletion that happen on the level of phonological form (Chomsky 1995:Ch. 3), the meaning differences cannot be explained by an analysis that deletes material.

In a further variant of the VP coordination analysis, we have a trace that is related to *this record*. This would be a right-node-raising analysis. Borsley (2005) showed that such analyses are problematic. Among the problematic examples that he discusses is the following pair (see also Bresnan 1974:615).

    (23) a. He tried to persuade and convince him.
       b. *He tried to persuade, but couldn't convince, him.

The second example is ungrammatical if *him* is not stressed. In contrast, 23a is well formed even with unstressed *him*. So, if 23a were an instance of right node raising, the contrast would be unexpected. Borsley therefore excludes a right-node-raising analysis.

The third possibility to analyze sentences like 21 assumes discontinuous constituents and uses material twice: the two VPs *knows this record* and *loves this record* are coordinated, with the first VP being discontinuous. (See Crysmann 2001, Beavers & Sag 2004 for such proposals in the framework of HPSG.) However, discontinuous constituents are not usually assumed in the minimalist framework (see for instance Kayne 1994:67). Furthermore, Abeillé (2006) showed that there is evidence for structures in which lexical elements are coordinated directly. This means that one needs analyses like the CG analysis discussed above, which would result in the problems with the specifier/complement status just discussed.

Furthermore, Abeillé pointed out that NP coordinations in head-final languages like Korean and Japanese present difficulties for Merge-based analyses. A Japanese example is given in 24.

---

[20] See also Bartsch & Vennemann 1972:102, Dowty 1979:143, den Besten 1983:104–5, Klein 1985:8–9, and Eisenberg 1994 for similar observations and criticism of similar proposals in earlier versions of transformational grammar.

(24) Robin-to   Kim
     Robin-and Kim
      'Kim and Robin'

In the first step *Robin* is merged with *to*. In a second step *Kim* is merged. Since *Kim* is a specifier, one would expect that *Kim* is serialized before the head as is the case for other specifiers in head-final languages.

Chomsky tries to get rid of the unary branching structures of standard X′ theory, which were needed to project lexical items like pronouns and determiners into full phrases, referring to work by Muysken (1982). Muysken used the binary features MIN and MAX to classify syntactic objects as minimal (words or word-like complex objects) or maximal (syntactic objects that stand for complete phrases). Such a feature system can be used to describe pronouns and determiners as [+MIN, +MAX]. Verbs like *give*, however, are classified as [+MIN, −MAX]. They have to project in order to reach the [+MAX] level. If specifiers and complements are required to be [+MAX], then determiners and pronouns fulfill this requirement without having to project from X^0 via X′ to the XP level.

In Chomsky's system, the MIN/MAX distinction is captured with respect to the completeness of heads (complete = phrase) and to the property of being a lexical item. However, there is a small but important difference between Muysken's and Chomsky's proposals: the predictions with regard to the coordination data that was discussed above. Within the category system of X′ theory, it is possible to combine two X^0s to get a new, complex X^0. This new object has basically the same syntactic properties that the simple X^0s have (see Jackendoff 1977:51 and Gazdar et al. 1985). In Muysken's system, the coordination rule (or the lexical item for the conjunction) can be formulated such that the coordination of two +MIN items is a +MIN item. In Chomsky's system, an analogous rule cannot be defined, since the coordination of two lexical items is not a lexical item any longer.

Like Chomsky in his recent minimalist work, CG (Ajdukiewicz 1935) and HPSG (Pollard & Sag 1987, 1994:39–40) do not (strictly) adhere to X′ theory. Both theories assign the symbol NP to pronouns (for CG, see Steedman & Baldridge 2006:615; see Steedman 2000:§4.4 for the incorporation of lexical type raising in order to accommodate quantification). The phrase *likes Mary* and the word *sleeps* have the same category in CG (s\np). In both theories it is not necessary to project a noun like *tree* from N^0 to $\bar{N}$ in order to be able to combine it with a determiner or an adjunct.

Determiners and monovalent verbs in controlled infinitives are not projected from an X^0 level to the XP level in many HPSG analyses, since the valence properties of the respective linguistic objects (an empty SUBCAT or COMPS list) are sufficient to determine their combinatoric potential and hence their distribution (Müller 1996b, 1999a). If the property of being minimal is needed for the description of a phenomenon, the binary feature LEX is used in HPSG (Pollard & Sag 1987:172, 1994:22). This feature is not needed, however, for the distinction between specifiers and complements. The distinction is governed by principles that map elements of an argument structure list (ARG-ST) onto valence lists that are the value of the SPECIFIER and the COMPLEMENTS feature (abbreviated as SPR and COMPS, respectively).[21]

Roughly speaking, the specifier in a verbal projection is the least oblique argument of the verb for configurational languages like English. Since the argument structure list is

---

[21] Some authors assume a three-way distinction between subjects, specifiers, and complements.

ordered according to the obliqueness hierarchy of Keenan & Comrie 1977, the first ele-
ment of this list is the least oblique argument of a verb, and this argument is mapped to
the SPR list. The element in the SPR list is realized to the left of the verb in SVO lan-
guages like English. The elements in the COMPS list are realized to the right of their
head. Approaches like the one by Ginzburg and Sag (2000:34, 364) that assume that
head-complement phrases combine a word with its arguments have the same problem
with coordinations like 21, since the head of the VP is not a word.[22] However, this re-
striction for the head can be replaced by one that refers to the LEX feature rather than to
the property of being a word or lexical item.

   Pollard and Sag and Ginzburg and Sag assume flat structures for English. Since one
of the daughters is marked as lexical, it follows that the rule does not combine a head
with a subset of its complements and then apply a second time to combine the result
with further complements. Therefore a structure like 25a is excluded, since *gave John* is
not a word and hence cannot be used as the head daughter in the rule.

   (25) a. [[gave John] a book]
         b. [gave John a book]

Instead of 25a, only analyses like 25b are admitted; that is, the head is combined with
all of its arguments all in one go. The alternative is to assume binary branching struc-
tures (Müller 2013e, Müller & Ørsnes 2013b:§1.2.2). In such an approach, the head-
complement schema does not restrict the word/phrase status of the head daughter. The
binary branching structures in HPSG correspond to external Merge in the MP.

   In the previous two sections we saw certain shortcomings of Chomsky's labeling def-
inition and problems with the coordination of lexical items. In the following section I
discuss Stabler's definition of Merge in minimalist grammar, which is explicit about la-
beling and in one version does not have the problems discussed above. I show that his
formalization corresponds rather directly to HPSG representations.

   2.3. MINIMALISM, CATEGORIAL GRAMMAR, AND HPSG. In §2.1 we saw that Chomsky's
papers leave many crucial details about labeling unspecified. Stabler's work is rela-
tively close to recent minimalist approaches, but is worked out with considerable preci-
sion (see also Stabler 2010:397, 399, 400 on formalization of post-GB approaches).
Stabler (2001) shows how Kayne's theory of remnant movement can be formalized and
implemented. Stabler refers to his particular way of formalizing minimalist theories as
MINIMALIST GRAMMARS (MG). Although his work can be regarded as a formalization of
Chomsky's minimalist ideas, Stabler's approach differs from Chomsky's in certain mat-
ters of detail. Stabler assumes that the results of the two Merge operations are not sets
but pairs. The head in a pair is marked by a pointer ('<' or '>'). Bracketed expressions
like {α, {α, β}} (discussed in §2.1) are replaced by trees like the one in 26.

   (26)

---
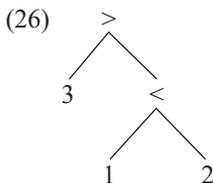
[22] As mentioned above, a multidomination approach with discontinuous constituents is a possible solution
for the analysis of 21 (see Crysmann 2001, Beavers & Sag 2004). But the coordination of lexical items has to
be possible in principle, as Abeillé (2006) has argued. Note also that the HPSG approach to coordination can-
not be taken over to the MP. The reason is that the HPSG proposals involve special grammar rules for coordi-
nation, and MP comes with the claim that there is only Merge. Hence the additional introduction of
combinatorial rules is not an option within the MP.

1 is the head in 26, 2 is the complement, and 3 the specifier. The pointer points to the part of the structure that contains the head. The daughters in a tree are ordered; that is, 3 is serialized before 1 and 1 before 2.

Stabler (2010:402) defines external Merge as follows in 27.

(27)

$$
em(t_1[=f], t_2[f]) = \begin{cases} \begin{array}{c} > \\ \diagup\diagdown \\ t_1 \quad\quad t_2 \end{array} \text{ if } t_1 \text{ has exactly 1 node} \\[2em] \begin{array}{c} > \\ \diagup\diagdown \\ t_2 \quad\quad t_1 \end{array} \text{ otherwise} \end{cases}
$$

=f is a selection feature and f the corresponding category. When $t_1[=f]$ and $t_2[f]$ are combined, the result is a tree in which the selection feature of t_1 and the respective category feature of t_2 are deleted. The upper tree in 27 represents the combination of a (lexical) head with its complement. t_1 is positioned before t_2. The condition that t_1 has to have exactly one node corresponds to Chomsky's assumption that the first Merge is a Merge with a complement and that all further applications of Merge are Merges with specifiers (Chomsky 2008:146).

A common assumption in current minimalist theories is that movement is feature driven: that is, an item moves to a special position in a tree structure to check a feature. An example would be a WH-element that is specified as −WH and has to check this feature in a specifier position of a CP. In Stabler's system, movement and feature checking are done in the definition of internal Merge, which is given in 28.[23]

(28) $im(t_1[+f]) = $

$$
\begin{array}{c} > \\ \diagup\diagdown \\ \overset{>}{t_2} \quad\quad t_1\{t_2[-f]^> \mapsto \varepsilon\} \end{array}
$$

t_1 is a tree with a subtree t_2 that has the feature f with the value '−'. This subtree is deleted ($t_2[-f]^> \mapsto \varepsilon$), and a copy of the deleted subtree without the −f feature ($\overset{>}{t_2}$) is positioned in specifier position. The element in specifier position has to be a maximal projection. This requirement is visualized by the raised '>'.

Stabler provides an example derivation for the clause in 29.

(29) who Marie praises

Praises is a two-place verb with two =D features. This encodes the selection of two determiner phrases. *Who* and *Marie* are two Ds, and they fill the object and subject positions of the verb. The resulting verbal projection *Marie praises who* is embedded under an empty complementizer that is specified as +WH and hence provides the position for the movement of *who*, which is placed in the specifier position of CP by the application of internal Merge. The −WH feature of *who* is deleted, and the result of the application of internal Merge is *who Marie praises*.

This analysis has a problem that was pointed out by Stabler himself in unpublished work cited by Veenstra (1998:124): it makes incorrect predictions in the case of monovalent verbs. If a verb is combined with an NP, the definition of external Merge in 27

---

[23] In addition to what is shown in 28, Stabler's definition contains a variant of the shortest move constraint, which is irrelevant for the discussion at hand and hence is omitted.

treats this NP as a complement and serializes it to the right of the head.[24] Instead of analyses of sentences like 30a, one gets analyses of strings like 30b.[25]

> (30) a. Max sleeps.
>   b. *Sleeps Max.

To solve this problem, Stabler assumes that monovalent verbs are combined with a nonovert object (Veenstra (1998:61, 124), quoting Stabler's unpublished work, also adopts this solution).[26] With such an empty object the resulting structure contains the empty object as a complement. The empty object is serialized to the right of the verb, and *Max* is the specifier and hence serialized to the left of the verb, as in 31.

> (31) Max sleeps __ .

Of course, any analysis of this kind is both stipulative and entirely ad hoc, being motivated only by the wish to have uniform structures. Moreover, it exemplifies precisely one of the methodological deficiencies of transformational generative grammar (the excessive appeal to uniformity) discussed at length by Culicover and Jackendoff (2005: §2.1.2).

Apart from the two operations that were defined in 27 and 28, there are no other operations in MG.[27] In addition to the problems with monovalent verbs, this results in the problem that was discussed in §2.2: there is no analysis with a direct combination of verbs for 21—repeated here as 32.

> (32) He [knows and loves] this record.

The reason is that the combination of *knows*, *and*, and *loves* consists of three nodes, and the Merge of *knows and loves* with *this record* would make *this record* the specifier of the structure. Therefore, *this record* would be serialized before *knows and loves*, contrary to the facts. Since the set of languages that can be generated with MGs contains the languages that can be generated with certain TAGs and with COMBINATORIAL CG (Michaelis 2001), the existence of a CG analysis implies that the coordination examples can be derived in MGs somehow. But for linguists, the fact that it is possible at all to generate a certain string (the weak capacity of a grammar) is of less significance. The actual structures that are licensed by the grammar are important (the strong capacity).

DIRECTIONAL MINIMALIST GRAMMARS AND CATEGORIAL GRAMMAR. In addition to the reintroduction of X^0 categories, this problem can be solved by changing the definition of Merge in a way that allows heads to specify the direction of combination with their arguments: Stabler (2011:635) suggests marking the position of an argument relative to its head together with the selection feature and gives the following redefinition of external Merge.

---

[24] Compare also Chomsky's definition of specifier and complement in §2.2.

[25] More elaborated analyses assume that the subject of a verb has to move to the specifier position of a Tense Phrase (TP). These analyses would not license 30b, but they would fail to derive 30a since *Max* would be treated as a complement rather than a specifier.

[26] An alternative that was pointed out to me by Joost Kremers would be to assume an empty verbal head that takes *sleeps* as complement and *Max* as subject. Such an analysis is often assumed for ditransitive verbs in minimalist theories that assume Larsonian verb shells (Larson 1988). Although this may work technically, there is no theory-external evidence for such an empty head. On the contrary: Müller and Wechsler (2014: §6.1, §7) pointed out empirical problems for analyses based on little v.

[27] For extensions see Frey & Gärtner 2002:§3.2.

(33)

$$
\text{em}(t_1[\alpha], t_2[x]) = \begin{cases} \overset{<}{\underset{t_1 \quad t_2}{\bigwedge}} & \text{if } \alpha \text{ is } =x \\[2em] \overset{>}{\underset{t_2 \quad t_1}{\bigwedge}} & \text{if } \alpha \text{ is } x= \end{cases}
$$

The position of the equals sign specifies on which side of the head an argument has to be realized. This corresponds to FORWARD and BACKWARD application in CG. Stabler calls this form of grammar DIRECTIONAL MG (DMG). This variant of MG avoids the problem with monovalent verbs, and the coordination data is unproblematic as well if one assumes that the conjunction is a head with a variable category that selects for elements of the same category to the left and to the right of itself. *Know* and *love* would both select an object to the right and a subject to the left, and this requirement would be transferred to the result of coordinating *knows* and *loves*. See Steedman 1991:264 for the details of the CG analysis and Bouma & van Noord 1998:52 for an earlier HPSG proposal involving directionality features along the lines suggested by Stabler for his DMGs.

MINIMALIST GRAMMARS AND HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR. The notation for marking the head of a structure with '>' and '<' corresponds directly to the HPSG representation of heads. Since HPSG is a sign-based theory, information about all relevant linguistic levels is represented in descriptions (phonology, morphology, syntax, semantics, information structure). An example is given in 34: the lexical entry for the word *grammar*.

(34)
$$
\begin{bmatrix} \text{PHON} & \langle \text{'gram}\ni\text{r} \rangle \\[1em] \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SPR} & \langle \text{DET} \rangle \\ \textit{cat} \end{bmatrix} \\[2em] \text{CONT} & \dots \begin{bmatrix} \text{INST } X \\ \textit{grammar} \end{bmatrix} \\[1em] \textit{loc} \end{bmatrix} \\[3em] \textit{word} \end{bmatrix}
$$

The part of speech of *grammar* is noun. In order to form a complete phrase, it requires a determiner. This is represented by giving the SPR feature the value ⟨DET⟩. Semantic information is listed under CONT. The details cannot be explained here further, but the interested reader is referred to Pollard & Sag 1994 or Müller 2013d.

Since we are dealing with syntactic aspects exclusively, only a subset of the used features is relevant: valency information and information about part of speech and certain morphosyntactic properties that are relevant for the external distribution of a phrase are represented in a feature description under the path SYNSEM|LOC|CAT. The features that are particularly interesting here are the so-called head features. Head features are shared between a lexical head and its maximal projection. The head features are located inside CAT and are grouped together under the path HEAD. Complex hierarchical structure is also modeled with feature value pairs. The constituents of a complex linguistic object are usually represented as parts of the representation of the complete object. For instance, there is a feature HEAD-DAUGHTER, the value of which is a feature structure that models a linguistic object that contains the head of a phrase. The head feature principle (35) refers to this daughter and ensures that the head features of the head daughter are identical with the head features of the mother node: that is, they are identical with the head features of the complete object.

(35) $\textit{headed-phrase} \Rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD} \;\boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|HEAD} \;\boxed{1} \end{bmatrix}$

Identity is represented by boxes with the same number.

Ginzburg and Sag (2000:30) represent all daughters of a linguistic object in a list that is given as the value of the DAUGHTERS attribute. The value of the feature HEAD-DAUGHTER is identified with one of the elements of the DAUGHTERS list.

(36) a. $\begin{bmatrix} \text{HEAD-DTR} \;\boxed{1} \\ \text{DTRS} \qquad \langle\, \boxed{1}\, \alpha,\, \beta \rangle \end{bmatrix}$

     b. $\begin{bmatrix} \text{HEAD-DTR} \;\boxed{1} \\ \text{DTRS} \qquad \langle\, \alpha,\, \boxed{1}\, \beta \rangle \end{bmatrix}$

α and β are shorthands for descriptions of linguistic objects. The important point about the two descriptions in 36 is that the head daughter is identical with one of the two daughters, which is indicated by the $\boxed{1}$ in front of α and β, respectively. In the first feature description, the first daughter is the head; in the second description, the second daughter is the head. Because of the head feature principle, the syntactic properties of the whole phrase are determined by the head daughter. That is, the syntactic properties of the head daughter correspond to the label in Chomsky's definition. This notation corresponds exactly to the one that is used by Stabler: 36a is equivalent to 37a, and 36b is equivalent to 37b.

(37) a.

            <
       /\
    α β

     b.

            >
       /\
    α β
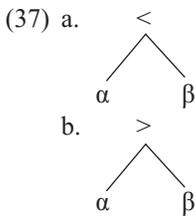
An alternative structuring of this basic information, discussed by Pollard and Sag (1994:Ch. 9), eliminates the DAUGHTERS feature, using instead the two features HEAD-DAUGHTER and NON-HEAD-DAUGHTERS. This gives rise to feature descriptions like 38a, which corresponds directly to Chomsky's set-based representations, discussed in §2.1 and repeated here as 38b.

(38) a. $\begin{bmatrix} \text{HEAD-DTR} \qquad\quad \alpha \\ \text{NON-HEAD-DTRS} \quad \langle \beta \rangle \end{bmatrix}$

     b. $\{\alpha, \{\alpha, \beta\}\}$

The representation in 38a does not contain information about the linear precedence of α and β. The linear precedence of constituents is constrained by linear precedence rules, which are represented independently from constraints regarding (immediate) dominance.

The definition of internal Merge in 28 corresponds to the head-filler schema in HPSG (Pollard & Sag 1994:164). Stabler's derivational rule deletes the subtree $t_2[-f\,]^{>}$. HPSG is monotonic; that is, nothing is deleted in structures that are licensed by a grammar. Instead of deleting t_2 inside of a larger structure, structures containing an empty element (NB—not a tree) are licensed directly.[28] Both in Stabler's definition and in the HPSG schema, t_2 is realized as a filler in the structure. In Stabler's definition of internal Merge the category of the head daughter is not mentioned, but Pollard and Sag (1994:164) re-

---

[28] See Bouma et al. 2001 for a traceless analysis of extraction in HPSG and Müller 2013c:§11.9 for a general discussion of empty elements.

strict the head daughter to be a finite verbal projection. Chomsky (2007:17) assumes that all operations but external Merge operate on phase level. Chomsky assumes that CP and v*P are phases. If this constraint is incorporated into the definition in 28, the restrictions on the label of t_1 would have to be extended accordingly. In HPSG, sentences like 39 have been treated as VPs, not as CPs, and hence Pollard and Sag's requirement that the head daughter in the head-filler schema be verbal corresponds to Chomsky's restriction.

 (39) Bagels, I like.

Hence, despite minor presentational differences, we may conclude that the formalization of internal Merge and that of the head-filler schema are very similar.

 An important difference between HPSG and Stabler's definition is that 'movement' is not feature driven in HPSG. This is an important advantage since feature-driven movement cannot deal with instances of so-called 'altruistic' movements (Fanselow 2003), that is, movements of a constituent that happen in order to make room for another constituent in a certain position. See also Bildhauer & Cook 2010:72 for a discussion of such altruistic multiple frontings in German.

 A further difference between general X′ theory and Stabler's formalization of internal Merge on the one hand and HPSG on the other is that in the latter case there is no restriction regarding the completeness (or valence 'saturation') of the filler daughter. Whether the filler daughter has to be a maximal projection (English) or not (German) follows from restrictions that are enforced locally when the trace is combined with its head. This makes it possible to analyze sentences like 40 without remnant movement.[29]

 (40) Gelesen$_i$ hat$_j$ das Buch keiner __ $_i$ __ $_j$
   read has the book nobody
    'Nobody has read the book.'

In contrast, Stabler is forced to assume an analysis like that in 41b (see also G. Müller 1998 for a remnant movement analysis). In a first step, *das Buch* is moved out of the VP (41a), and in a second step the emptied VP is fronted, as in 41b.

 (41) a. Hat [das Buch]$_j$ [keiner [$_{VP}$ __ $_j$ gelesen]].
   b. [$_{VP}$ __ $_j$ Gelesen]$_i$ hat [das Buch]$_j$ [keiner __ $_i$].

Haider (1993:281), De Kuthy and Meurers (2001:§2), and Fanselow (2002) showed that such remnant movement analyses are problematic for German. The only phenomenon that Fanselow identified as requiring a remnant movement analysis is the problem of multiple frontings (see Müller 2003a for an extensive discussion of relevant data). An alternative analysis of these multiple frontings that uses an empty verbal head in the *Vorfeld* (the position before the finite verb in root clauses) is developed in Müller 2005a,b, 2013b, but it does not assume that adjuncts or arguments like *das Buch* in 41b are extracted from the *Vorfeld* constituent. Instead of the remnant movement analysis, the mechanism of argument composition from CG (Geach 1970, Hinrichs & Nakazawa 1994) is used to ensure the proper realization of arguments in the sentence. Chomsky (2007:20) already uses argument composition as part of his analysis of TPs and CPs. Hence, both remnant movement and argument composition are already assumed in recent minimalist proposals. The HPSG alternative, however, would appear to need less theoretical apparatus.

---

[29] See also Müller & Ørsnes 2013a for an analysis of object shift in Danish that can account for verb fronting without remnant movement. The analysis does not have any of the problems that remnant movement analyses have.

Finally, it should be mentioned that all transformational accounts have problems with across-the-board extractions like 7b and 42 in which one element corresponds to several gaps.

(42) The man who$_i$ [Mary loves __ $_i$] and [Sally hates __ $_i$] computed my tax.

This problem was solved for GPSG by Gazdar (1981), and the solution carries over to HPSG. Furthermore, movement paradoxes (Bresnan 2001:Ch. 2) can be avoided by not sharing all information between filler and gap, a solution that is not available for transformational accounts, which usually assume identity of filler and gap or—as in the copy theory of movement—assume that a derivation contains multiple copies of one object, only one of which is spelled out. See also Borsley 2012 for further puzzles for, and problems of, movement-based approaches.

3. CONCLUSIONS. In §1 I showed that Jackendoff's proposal to store certain constructions together with phrase-structural information is problematic. In particular, all phenomena that interact with valence-changing processes like active/passive alternations and/or with derivational morphology should not be treated in terms of complete tree structures that are assigned a meaning. Rather, lexical items should represent the combinatoric potential of the involved elements. I agree with all other things that were said in his discussion note in *Language* (Jackendoff 2011). In particular, there are various reasons to model grammar in a constraint-based fashion, as is done in LFG (Kaplan 1995), HPSG (King 1999, Pollard 1999, Richter 2007), and also variants of minimalism (Veenstra 1998). I have also demonstrated various shortcomings of extant definitions of Merge (labeling and determination of specifier/complement status) and shown how Merge can be reformulated using feature-value pairs. This shows that a rather straightforward constraint-based representation of Merge is possible. If all shortcomings of the existing proposals are fixed, one ends up with a representation that has been standardly assumed in the HPSG community for quite some time now. However, Sag (2010) has argued that the head-filler schema should be viewed as a most general common denominator for a 'family' of filler-gap constructions in English, each of which involves more specific constraints regarding meaning (interrogative, exclamative, etc.) and form (inversion, filler category, and so forth). In addition, Bildhauer and Cook (2010), among others, have suggested information-structural constraints on head-filler structures of various kinds.

So, there is a constraint-based formalization of Merge that may play a role in an adequate theory of language. However, as is repeatedly pointed out by researchers working in constructionist theories: Merge alone is not sufficient for licensing structures if one does not want to make stipulative assumptions. An example is Jackendoff's N-P-N construction (Jackendoff 2008). As Jackendoff argued, there is no element in the N-P-N construction that could plausibly function as a head. Since it is assumed in minimalism that all structures are headed and licensed by binary Merge, Jackendoff's N-P-N construction and similar constructions pose a challenge for minimalist theories. What proportion of a language has to be described with phrasal constructions in Jackendoff's sense and what has to be described lexically is an open issue that has to be decided empirically. See for instance Budde 2010, Maché 2010, G. Müller 2011, and Ørsnes 2011 for proposals for analyses of Jacobs's verbless directives (2008) (an example for such directives is *In den Müll mit diesen Klamotten!* 'Into the garbage with these duds!').

To conclude: A theory of language has to meet at least the following desiderata (many of which were mentioned in or follow from Jackendoff's discussion note):

• constraint-based formalization (Pullum & Scholz 2001, Pullum 2007, Sag & Wasow 2011)

- strongly lexicalist orientation (Sag & Wasow 2011 and §1 of this discussion note)
- parallel/sign-based architecture, including constraints on phonology, morphology, syntax, semantics, and information structure and the interactions between the various levels of linguistic description (Kuhn 2007, Jackendoff 2011)
- inclusion of both headed and nonheaded configurations (Jackendoff 2008, Jacobs 2008)
- possibility of describing complex linguistic objects rather than just lexical items (Sag 1997, Kay & Fillmore 1999, Jackendoff 2008, Jacobs 2008)

Sag and Wasow (2011) added another important property: that of being surface-oriented. Being surface-oriented does not mean that there is no movement or anything equivalent. Nor does it preclude the existence of phonetically empty elements. (In fact, many proposals with empty elements can be recast without empty elements (Bar-Hillel et al. 1961), but crucial insights sometimes get lost in grammars that do not use such elements (Bender 2000, Sag et al. 2003:464, Müller 2004).) What it does mean is that analyses like the one by Laenzlinger (2004) are excluded. Laenzlinger assumes a Kayneian analysis of German: that is, the base order is assumed to be SVO (contrary to all wisdom in the linguistic community that treats German as an SOV language (Bach 1962, Bierwisch 1963, Reis 1974, Koster 1975, Thiersch 1978, Haider 2010)), and the visible orders are derived from this base configuration by a set of iterated movements that can only be established because several empty functional heads are introduced into the theory. In addition to being psycholinguistically implausible, it is unclear how such grammars could be acquired in language acquisition, given what we know about innateness and language acquisition today (see below).

Returning to the desiderata for linguistic theories, HPSG is a framework that has the desired properties. Most of them were available in early versions of HPSG, and the possibility of describing complex linguistic objects via multiple inheritance hierarchies was established in 1997 (Sag 1997). Since then, HPSG has been well suited to meet the requirements of current linguistic research. HPSG differs from many minimalist proposals in assuming rich descriptions of the arguments of a head: the linguistic objects are described by complex, internally structured feature descriptions. Note, however, that it has been argued in the MP literature that a description of this kind, based on the notion of feature values and feature structures, is also needed in the minimalist framework (Sauerland & Elbourne 2002:290–91).

I have shown in §1 that instead of a phrasal configuration for *rumble* constructions, a rich lexical representation (a lexical construction in the sense of Kay 2005) is needed, which is used in linguistic structures that are licensed by the head-specifier schema, the head-complement schema, and the head-filler schema (equivalent to the subcases of Merge and Move). Therefore, there seems to be convergence between theories in several respects. Of course, this is an impression that one gets only if one ignores certain details and also sociological issues. For instance, the question of how detailed a linguistic analysis should be is answered differently in different theoretical communities: see, for instance, Borsley 2006 and Sag 2010 for the fine-grained distinctions that set apart various types of filler-gap constructions in English, which are usually not addressed in GB/MP proposals. Another question is what counts as an explanation and what primitives/assumptions come without additional costs. For instance, Cinque and Rizzi (2010: 54, 61) assume that every morphosyntactic property corresponds to an independent syntactic head. Furthermore, they assume (pp. 55, 57) that there are at least 400 of these heads and that the respective categories are innately specified as part of UG. Structures have to contain these heads, but the heads can be empty. This is in sharp contrast to other theories that try to avoid unnecessary structure and the stipulation of empty heads.

Since identifiability in the limit, in the sense of Gold 1967, turned out to be irrelevant for cognitive science (Pullum 2003, Johnson 2004), and hence there is no mathematical proof for the poverty of the stimulus, and given that it could be shown that pace Chomsky 1957:16 statistical information in the linguistic signal is very rich (Abney 1996: §4.2) and can be used to infer complex structure (for instance about auxiliary inversion in English, which played an important role in the poverty-of-the-stimulus discussion (Bod 2009)), it seems rather unlikely that humans have an innate, linguistically rich UG.[30] This has been acknowledged by many working in different paradigms and will influence the way in which analyses are judged. For instance, analyses that motivate an empty functional Topic head on the basis of an overt morpheme in Japanese, or empty AgrO heads in German on the basis of object agreement in Basque, or an empty anti-passive head in German on the basis of Chukchi can no longer be accepted as well motivated.

In addition to the new results regarding the poverty of the stimulus, we now have new experimental methods and huge amounts of data in the form of unannotated and annotated corpora (see for instance Schäfer & Bildhauer 2012) that can be used to verify claims and explore phenomena in greater depth (Müller 2007b, Meurers & Müller 2009).

Though there are huge differences in details of the analyses and in the way research is done in different communities, I am not as pessimistic as Sternefeld and Richter (2012), since there are also some signs of convergence, as I hope to have shown in this discussion note. Perhaps there will be further convergence in the next years, and we can unify everything.

REFERENCES

ABBOTT, BARBARA. 1976. Right node raising as a test for constituenthood. *Linguistic Inquiry* 7.4.639–42.

ABEILLÉ, ANNE. 2006. In defense of lexical coordination. *Empirical Issues in Formal Syntax and Semantics* 6.7–36. Online: http://www.cssp.cnrs.fr/eiss6/.

ABNEY, STEVEN P. 1996. Statistical methods and linguistics. *The balancing act: Combining symbolic and statistical approaches to language* (Language, speech, and communication), ed. by Judith L. Klavans and Philip Resnik, 1–26. Cambridge, MA: MIT Press.

AJDUKIEWICZ, KASIMIR. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1.1–27.

ASUDEH, ASH; MARY DALRYMPLE; and IDA TOIVONEN. 2008. Constructions with lexical integrity: Templates as the lexicon-syntax interface. *Proceedings of the LFG '08 Conference*, 68–88. Online: http://cslipublications.stanford.edu/LFG/13.

ASUDEH, ASH; MARY DALRYMPLE; and IDA TOIVONEN. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1.1.1–54.

BACH, EMMON. 1962. The order of elements in a transformational grammar of German. *Language* 8.3.263–69.

BAR-HILLEL, YEHOSHUA; MICHA A. PERLES; and ELIAHU SHAMIR. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 14.2.143–72.

BARTSCH, RENATE, and THEO VENNEMANN. 1972. *Semantic structures: A study in the relation between semantics and syntax.* (Athenäum-Skripten Linguistik 9.) Frankfurt: Athenäum.

BAUSEWEIN, KARIN. 1990. Haben kopflose Relativsätze tatsächlich keine Köpfe? *Strukturen und Merkmale syntaktischer Kategorien* (Studien zur deutschen Grammatik 39), ed. by Gisbert Fanselow and Sascha W. Felix, 144–58. Tübingen: Gunter Narr.

BEAVERS, JOHN, and IVAN A. SAG. 2004. Coordinate ellipsis and apparent non-constituent co-ordination. *Proceedings of the 11th International Conference on Head-Driven Phrase*

[30] Berwick et al. 2011 discusses challenges to Chomsky's poverty-of-the-stimulus argument, but ignores Bod's work although Bod's article appeared in the same journal and was published two years before Berwick and colleagues submitted the revised version of their paper. See also Chomsky 2013:39.

Structure Grammar (HPSG 2004), 48–69. Online: http://cslipublications.stanford.edu /HPSG/5/.

BENDER, EMILY M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford, CA: Stanford University dissertation. Online: http://faculty .washington.edu/ebender/dissertation/.

BENDER, EMILY M.; DANIEL P. FLICKINGER; and STEPHAN OEPEN. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, Taipei, Taiwan, ed. by John Carroll, Nelleke Oostdijk, and Richard Sutcliffe, 8–14.

BERWICK, ROBERT C., and SAMUEL DAVID EPSTEIN. 1995. On the convergence of the 'minimalist' syntax and categorial grammar. *Algebraic methods in language processing*, ed. by Anton Nijholt, Giuseppe Scollo, and Rene Steetskamp, 143–48. Enschede: University of Twente. Online: http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf.

BERWICK, ROBERT C.; PAUL PIETROSKI; BERACAH YANKAMA; and NOAM CHOMSKY. 2011. Poverty of the stimulus revisited. *Cognitive Science* 35.7.1207–42.

BIERWISCH, MANFRED. 1963. *Grammatik des deutschen Verbs*. (Studia grammatica 2.) Berlin: Akademie.

BILDHAUER, FELIX, and PHILIPPA COOK. 2010. German multiple fronting and expected topic-hood. *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2010)*, 68–79. Online: http://cslipublications.stanford.edu /HPSG/2010/.

BLEVINS, JAMES P. 2003. Passives and impersonals. *Journal of Linguistics* 39.3.473–520.

BOD, RENS. 2009. From exemplar to grammar: Integrating analogy and probability in language learning. *Cognitive Science* 33.4.752–93.

BORER, HAGIT. 2005. *Structuring sense, vol. 1: In name only*. Oxford: Oxford University Press.

BORSLEY, ROBERT D. 2005. Against ConjP. *Lingua* 115.4.461–82.

BORSLEY, ROBERT D. 2006. *Syntactic and lexical approaches to unbounded dependencies*. (Essex research reports in linguistics 49.) Essex: University of Essex. Online: http:// repository.essex.ac.uk/127/.

BORSLEY, ROBERT D. 2012. Don't Move! *Iberia: An International Journal of Theoretical Linguistics* 4.1.110–39.

BOUMA, GOSSE; ROBERT MALOUF; and IVAN A. SAG. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19.1.1–65.

BOUMA, GOSSE, and GERTJAN VAN NOORD. 1998. Word order constraints on verb clusters in German and Dutch. *Syntax and semantics, vol. 30: Complex predicates in nonderivational syntax*, ed. by Erhard W. Hinrichs, Andreas Kathol, and Tsuneko Nakazawa, 43–72. San Diego: Academic Press.

BRESNAN, JOAN. 1974. The position of certain clause-particles in phrase structure. *Linguistic Inquiry* 5.4.614–19.

BRESNAN, JOAN. 1982. The passive in lexical theory. *The mental representation of grammatical relations* (Cognitive theory and mental representation), ed. by Joan Bresnan, 3–86. Cambridge, MA: MIT Press.

BRESNAN, JOAN. 2001. *Lexical-functional syntax*. Oxford: Blackwell.

BRESNAN, JOAN, and JANE GRIMSHAW. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9.3.331–92.

BRESNAN, JOAN, and SAM A. MCHOMBO. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13.181–254.

BRISCOE, TED J., and ANN COPESTAKE. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25.4.487–526. Online: http://acl.ldc.upenn.edu/J/J99/J99 -4002.pdf.

BUDDE, MONIKA. 2010. Konstruktionen integrativ: Topik-Konstruktionen als reinsyntaktisches Pendant zu den lexikalisch verankerten Komplement-Konstruktionen. Paper presented at Konstruktionsgrammatik: Neue Perspektiven zur Untersuchung des Deutschen und Englischen: Internationale Fachtagung an der Christian-Albrechts-Universität zu Kiel vom 18. bis 20. Februar 2010.

CANDITO, MARIE-HÉLÈNE. 1996. A principle-based hierarchical representation of LTAGs. *Proceedings of the International Conference on Computational Linguistics (COLING '96)* 16.194–99.

CHOMSKY, NOAM. 1957. *Syntactic structures*. The Hague: Mouton.

CHOMSKY, NOAM. 1981. *Lectures on government and binding*. Dordrecht: Foris.

CHOMSKY, NOAM. 1995. *The minimalist program*. (Current studies in linguistics 28.) Cambridge, MA: MIT Press.

CHOMSKY, NOAM. 2005. Three factors in language design. *Linguistic Inquiry* 36.1.1–22.

CHOMSKY, NOAM. 2007. Approaching UG from below. *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics* (Studies in generative grammar 89), ed. by Uli Sauerland and Hans-Martin Gärtner, 1–29. Berlin: Mouton de Gruyter.

CHOMSKY, NOAM. 2008. On phases. *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–66. Cambridge, MA: MIT Press.

CHOMSKY, NOAM. 2013. Problems of projection. *Lingua* 130.33–49.

CINQUE, GUGLIELMO, and LUIGI RIZZI. 2010. The cartography of syntactic structures. *The Oxford handbook of linguistic analysis*, ed. by Bernd Heine and Heiko Narrog, 51–65. Oxford: Oxford University Press.

CITKO, BARBARA. 2008. Missing labels. *Lingua* 118.7.907–44.

CLÉMENT, LIONEL, and ALEXANDRA KINYON. 2003. Generating parallel multilingual LFG-TAG grammars from a metagrammar. *Proceedings of the 41st annual meeting of the Association for Computational Linguistics (ACL 2003)*, 184–91.

CRYSMANN, BERTHOLD. 2001. Clitics and coordination in linear structure. *Clitics in phonology, morphology and syntax* (Linguistik aktuell/Linguistics today 36), ed. by Birgit Gerlach and Janet Grijzenhout, 121–59. Amsterdam: John Benjamins.

CULICOVER, PETER W., and RAY S. JACKENDOFF. 2005. *Simpler syntax*. Oxford: Oxford University Press.

DAVIS, ANTHONY R., and JEAN-PIERRE KOENIG. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76.1.56–91.

DE KUTHY, KORDULA, and WALT DETMAR MEURERS. 2001. On partial constituent fronting in German. *Journal of Comparative Germanic Linguistics* 3.3.143–205.

DEN BESTEN, HANS. 1983. On the interaction of root transformations and lexical deletive rules. *On the formal syntax of the Westgermania: Papers from the 3rd Groningen Grammar Talks, Groningen, January 1981* (Linguistik aktuell/Linguistics today 3), ed. by Werner Abraham, 47–131. Amsterdam: John Benjamins.

DONATI, C. 2006. On WH-head-movement. *WH-movement: Moving on* (Current studies in linguistics 42), ed. by Lisa Lai-Shen Cheng and Norbert Corver, 21–46. Cambridge, MA: MIT Press.

DOWTY, DAVID R. 1978. Governed transformations as lexical rules in a Montague grammar. *Linguistic Inquiry* 9.3.393–426.

DOWTY, DAVID R. 1979. *Word meaning and Montague grammar*. (Synthese language library 7.) Dordrecht: Reidel.

DOWTY, DAVID R. 2003. The dual analysis of adjuncts/complements in categorial grammar. *Modifying adjuncts* (Interface explorations 4), ed. by Ewald Lang, Claudia Maienborn, and Cathrine Fabricius-Hansen, 33–66. Berlin: Mouton de Gruyter.

EGG, MARKUS. 1999. Derivation and resolution of ambiguities in *wieder*-sentences. *Proceedings of the 12th Amsterdam Colloquium*, ed. by Paul J. E. Dekker, 109–14.

EISENBERG, PETER. 1994. *Grundriß der deutschen Grammatik*. 3rd edn. Stuttgart: J. B. Metzler.

FANSELOW, GISBERT. 2002. Against remnant VP-movement. *Dimensions of movement: From features to remnants* (Linguistik aktuell/Linguistics today 48), ed. by Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner, 91–127. Amsterdam: John Benjamins.

FANSELOW, GISBERT. 2003. Free constituent order: A minimalist interface account. *Folia Linguistica* 37.1–2.191–231.

FREUDENTHAL, DANIEL; JULIAN M. PINE; JAVIER AGUADO-OREA; and FERNAND GOBET. 2007. Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* 31.2.311–41.

FREUDENTHAL, DANIEL; JULIAN M. PINE; and FERNAND GOBET. 2006. Modeling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science* 30.2.277–310.

FREY, WERNER, and HANS-MARTIN GÄRTNER. 2002. On the treatment of scrambling and adjunction in minimalist grammars. *Proceedings of Formal Grammar 2002*, Trento, 41–52.

GAZDAR, GERALD. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12.2.155–84.

GAZDAR, GERALD; EWAN KLEIN; GEOFFREY K. PULLUM; and IVAN A. SAG. 1985. *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

GEACH, PETER THOMAS. 1970. A program for syntax. *Synthese* 22.3–17.

GINZBURG, JONATHAN, and IVAN A. SAG. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives*. (CSLI lecture notes 123.) Stanford, CA: CSLI Publications.

GOLD, MARK E. 1967. Language identification in the limit. *Information and Control* 10.5. 447–74.

GOLDBERG, ADELE E. 1995. *Constructions: A construction grammar approach to argument structure*. (Cognitive theory of language and culture.) Chicago: University of Chicago Press.

GOLDBERG, ADELE E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

GOLDBERG, ADELE E. 2013. Argument structure constructions versus lexical rules or derivational verb templates. *Mind & Language* 28.4.435–65.

GOLDBERG, ADELE E., and RAY S. JACKENDOFF. 2004. The English resultative as a family of constructions. *Language* 80.3.532–68.

GROSS, ANNEKE, and HENK VAN RIEMSDIJK. 1981. Matching effects in free relatives: A parameter of core grammar. *Theory of markedness in generative grammar: Proceedings of the 1979 GLOW Conference*, ed. by Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 171–216. Pisa: Scuola Normale Superiore.

GROß, THOMAS M., and TIMOTHY OSBORNE. 2009. Toward a practical dependency grammar theory of discontinuities. *SKY Journal of Linguistics* 22.43–90.

GUNJI, TAKAO. 1986. Subcategorization and word order. *Papers from the Second International Workshop on Japanese Syntax*, ed. by William J. Poser, 1–21. Stanford, CA: CSLI Publications.

HAIDER, HUBERT. 1993. *Deutsche Syntax—generativ: Vorstudien zur Theorie einer projektiven Grammatik*. (Tübinger Beiträge zur Linguistik 325.) Tübingen: Gunter Narr.

HAIDER, HUBERT. 2010. *The syntax of German*. (Cambridge syntax guides.) Cambridge: Cambridge University Press.

HAUGEREID, PETTER. 2009. *Phrasal subconstructions: A constructionalist grammar design, exemplified with Norwegian and English*. Trondheim: Norwegian University of Science and Technology.

HINRICHS, ERHARD W., and TSUNEKO NAKAZAWA. 1994. Linearizing AUXs in German verbal complexes. *German in head-driven phrase structure grammar* (CSLI lecture notes 46), ed. by John Nerbonne, Klaus Netter, and Carl J. Pollard, 11–38. Stanford, CA: CSLI Publications.

JACKENDOFF, RAY S. 1975. Morphological and semantic regularities in the lexicon. *Language* 51.3.639–71.

JACKENDOFF, RAY S. 1977. $\overline{X}$ *syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

JACKENDOFF, RAY S. 2008. Construction after construction and its theoretical challenges. *Language* 84.1.8–28.

JACKENDOFF, RAY S. 2011. What is the human language faculty? Two views. *Language* 87.3.586–624.

JACOBS, JOACHIM. 2008. Wozu Konstruktionen? *Linguistische Berichte* 213.3–44.

JACOBS, JOACHIM. 2009. Valenzbindung oder Konstruktionsbindung? Eine Grundfrage der Grammatiktheorie. *Zeitschrift für Germanistische Linguistik* 37.3.490–513.

JOHNSON, KENT. 2004. Gold's theorem and cognitive science. *Philosophy of Science* 71.4.571–92.

KALLMEYER, LAURA, and RAINER OSSWALD. 2012. A frame-based semantics of the dative alternation in lexicalized tree adjoining grammars. *Empirical issues in syntax and semantics*, vol. 9, ed. by Christopher Piñón, 167–84. Paris: CNRS Éditions.

KAPLAN, RONALD M. 1995. The formal architecture of lexical-functional grammar. *Formal issues in lexical-functional grammar* (CSLI lecture notes 47), ed. by Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, 7–27. Stanford, CA: CSLI Publications.

KAY, PAUL. 2005. Argument structure constructions and the argument-adjunct distinction. *Grammatical constructions: Back to the roots* (Constructional approaches to language 4), ed. by Mirjam Fried and Hans C. Boas, 71–98. Amsterdam: John Benjamins.

KAY, PAUL, and CHARLES J. FILLMORE. 1999. Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. *Language* 75.1.1–33.

KAYNE, RICHARD S. 1994. *The antisymmetry of syntax*. (Linguistic inquiry monographs 25.) Cambridge, MA: MIT Press.

KAYNE, RICHARD S. 2011. Why are there no directionality parameters? *West Coast Conference on Formal Linguistics (WCCFL)* 28.1–23.

KEENAN, EDWARD L., and BERNARD COMRIE. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.1.63–99.

KING, PAUL. 1999. Towards truth in head-driven phrase structure grammar. *Tübingen studies in head-driven phrase structure grammar* (Arbeitsberichte des SFB 340, 132), ed. by Valia Kordoni, 301–52. Tübingen: Universität Tübingen. Online: http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/132/132abs.html.

KISS, TIBOR. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum*. (Linguistische Arbeiten 333.) Tübingen: Max Niemeyer.

KLEIN, WOLFGANG. 1985. Ellipse, Fokusgliederung und thematischer Stand. *Ellipsen und fragmentarische Ausdrücke*, ed. by Reinhard Meyer-Hermann and Hannes Rieser, 1–24. Tübingen: Max Niemeyer.

KOENIG, JEAN-PIERRE. 1999. *Lexical relations*. (Stanford monographs in linguistics.) Stanford, CA: CSLI Publications.

KORDONI, VALIA. 2001. Linking experiencer-subject psych verb constructions in Modern Greek. *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2000)*, 198–213. Online: http://cslipublications.stanford.edu/HPSG/1/.

KOSTER, JAN. 1975. Dutch as an SOV Language. *Linguistic Analysis* 1.2.111–36.

KUHN, JONAS. 2007. Interfaces in constraint-based theories of grammar. *The Oxford handbook of linguistic interfaces*, ed. by Gillian Ramchand and Charles Reiss, 613–50. Oxford: Oxford University Press.

LABELLE, MARIE. 2007. Biolinguistics, the minimalist program, and psycholinguistic reality. *Snippets* 14. Online: http://www.ledonline.it/snippets/.

LAENZLINGER, CHRISTOPH. 2004. A feature-based theory of adverb syntax. *Adverbials: The interplay between meaning, context, and syntactic structure* (Linguistik aktuell/Linguistics today 70), ed. by Jennifer R. Austin, Stefan Engelberg, and Gisa Rauh, 205–52. Amsterdam: John Benjamins.

LARSON, RICHARD K. 1988. On the double object construction. *Linguistic Inquiry* 19.3.335–91.

LÜDELING, ANKE. 2001. *On particle verbs and similar constructions in German*. (Dissertations in linguistics.) Stanford, CA: CSLI Publications.

MACHÉ, JAKOB. 2010. Towards a compositional analysis of verbless directives in German. Paper presented at the 17th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2010).

MARANTZ, ALEC. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics (Proceedings of the 21st annual Penn Linguistics Colloquium)* 4.2.201–25. Online: http://www.ling.upenn.edu/papers/v4.2-contents.html.

MEURERS, WALT DETMAR. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24.2.161–217.

MEURERS, WALT DETMAR, and STEFAN MÜLLER. 2009. Corpora and syntax. *Corpus linguistics: An international handbook* (Handbooks of linguistics and communication science 29.2), ed. by Anke Lüdeling and Merja Kytö, 920–33. Berlin: Mouton de Gruyter.

MICHAELIS, JENS. 2001. *On formal properties of minimalist grammars*. Potsdam: Universität Potsdam dissertation.

MICHAELIS, LAURA A., and JOSEF RUPPENHOFER. 2001. *Beyond alternations: A constructional model of the German applicative pattern*. (Stanford monographs in linguistics.) Stanford, CA: CSLI Publications.

MÜLLER, GEREON. 1998. Incomplete category fronting. *A derivational approach to remnant movement in German*. (Studies in natural language and linguistic theory 42.) Dordrecht: Kluwer.

MÜLLER, GEREON. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. *Sprachliches Wissen zwischen Lexikon und Grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), ed. by Stefan Engelberg, Anke Holler, and Kristel Proost, 211–49. Berlin: De Gruyter.

MÜLLER, STEFAN. 1996a. The Babel-system—An HPSG fragment for German, a parser, and a dialogue component. *Proceedings of the fourth International Conference on the Practical Application of Prolog*, 263–77. London. Online: http://hpsg.fu-berlin.de/~stefan/Pub/babel.html.

MÜLLER, STEFAN. 1996b. Yet another paper about partial verb phrase fronting in German. *Proceedings of the International Conference on Computational Linguistics (COLING '96)* 16.800–805. Online: http://hpsg.fu-berlin.de/~stefan/Pub/pvp.html.

MÜLLER, STEFAN. 1999a. Deutsche Syntax deklarativ. *Head-Driven Phrase Structure Grammar für das Deutsche*. (Linguistische Arbeiten 394.) Tübingen: Max Niemeyer.

MÜLLER, STEFAN. 1999b. An HPSG-analysis for free relative clauses in German. *Grammars* 2.1.53–105.

MÜLLER, STEFAN. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German*. (Studies in constraint-based lexicalism 13.) Stanford, CA: CSLI Publications.

MÜLLER, STEFAN. 2003a. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31.1.29–62.

MÜLLER, STEFAN. 2003b. Object-to-subject-raising and lexical rule: An analysis of the German passive. *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2003)*, 278–97. Online: http://cslipublications.stanford.edu/HPSG/2003/.

MÜLLER, STEFAN. 2003c. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39.2.275–325.

MÜLLER, STEFAN. 2004. Elliptical constructions, multiple frontings, and surface-based syntax. *Proceedings of Formal Grammar 2004*, Nancy, 91–106. Online: http://hpsg.fu-berlin.de/~stefan/Pub/surface.html.

MÜLLER, STEFAN. 2005a. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201.3–39.

MÜLLER, STEFAN. 2005b. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203.297–330.

MÜLLER, STEFAN. 2006. Phrasal or lexical constructions? *Language* 82.4.850–83.

MÜLLER, STEFAN. 2007a. Phrasal or lexical constructions: Some comments on underspecification of constituent order, compositionality, and control. *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar (HPSG 2006)*, 373–93. Online: http://cslipublications.stanford.edu/HPSG/8/.

MÜLLER, STEFAN. 2007b. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. *Sprachkorpora—Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), ed. by Gisela Zifonun and Werner Kallmeyer, 70–90. Berlin: Walter de Gruyter. Online: http://hpsg.fu-berlin.de/~stefan/Pub/intro-korpus.html.

MÜLLER, STEFAN. 2013a. The CoreGram project: Theoretical linguistics, theory development and verification. Berlin: Freie Universität Berlin, MS. Online: http://hpsg.fu-berlin.de/~stefan/Pub/coregram.html.

MÜLLER, STEFAN. 2013b. German sentence structure: An analysis with special consideration of so-called multiple fronting. *Empirically oriented theoretical morphology and syntax*. Berlin: Language Science Press, to appear. Online: http://hpsg.fu-berlin.de/~stefan/Pub/gs.html.

MÜLLER, STEFAN. 2013c. *Grammatiktheorie*. 2nd edn. (Stauffenburg Einführungen 20.) Tübingen: Stauffenburg. Online: http://hpsg.fu-berlin.de/~stefan/Pub/grammatiktheorie.html.

MÜLLER, STEFAN. 2013d. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3rd edn. (Stauffenburg Einführungen 17.) Tübingen: Stauffenburg. Online: http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html.

MÜLLER, STEFAN. 2013e. HPSG—A synopsis. *Syntax: Ein internationales Handbuch zeitgenössischer Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft),

2nd edn., ed. by Artemis Alexiadou and Tibor Kiss. Berlin: Walter de Gruyter, to appear.

MÜLLER, STEFAN, and BJARNE ØRSNES. 2013a. Towards an HPSG analysis of object shift in Danish. *Formal Grammar: 17th and 18th international conferences: FG 2012/2013, Opole, Poland, August 2012, revised selected papers; Düsseldorf, Germany, August 2013, proceedings*, 69–89. Online: http://hpsg.fu-berlin.de/~stefan/Pub/object-shift.html.

MÜLLER, STEFAN, and BJARNE ØRSNES. 2013b. *Danish in head-driven phrase structure grammar*. (Empirically oriented theoretical morphology and syntax.) Berlin: Language Science Press, to appear. Online: http://hpsg.fu-berlin.de/~stefan/Pub/danish.html.

MÜLLER, STEFAN, and STEPHEN MARK WECHSLER. 2014. Lexical approaches to argument structure. *Theoretical Linguistics*, to appear. Online: http://hpsg.fu-berlin.de/~stefan/Pub/arg-st.html.

MUYSKEN, PETER. 1982. Parameterizing the notion of 'head'. *Journal of Linguistic Research* 2.57–75.

NOONAN, MICHAEL. 1994. A tale of two passives in Irish. *Voice: Form and function* (Typological studies in language 27), ed. by Barbara Fox and Paul J. Hopper, 279–311. Amsterdam: John Benjamins.

ØRSNES, BJARNE. 2011. „Wohin mit Konstruktionen?"—The syntax and semantics of non-verbal questions in German and Danish. Berlin: Freie Universität Berlin, MS.

ÖZKARAGÖZ, İNCI. 1986. Monoclausal double passives in Turkish. *Studies in Turkish linguistics* (Typological studies in language 8), ed. by Dan I. Slobin and Karl Zimmer, 77–92. Amsterdam: John Benjamins.

OSBORNE, TIMOTHY; MICHAEL PUTNAM; and THOMAS M. GROSS. 2011. Bare phrase structure, label-less trees, and specifier-less syntax: Is minimalism becoming a dependency grammar? *The Linguistic Review* 28.3.315–64.

OTT, DENNIS. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42.1.183–92.

POLLARD, CARL J. 1999. Strong generative capacity in HPSG. *Lexical and constructional aspects of linguistic explanation* (Studies in constraint-based lexicalism 1), ed. by Gert Webelhuth, Jean-Pierre Koenig, and Andreas Kathol, 281–98. Stanford, CA: CSLI Publications.

POLLARD, CARL J., and IVAN A. SAG. 1987. *Information-based syntax and semantics*. (CSLI lecture notes 13.) Stanford, CA: CSLI Publications.

POLLARD, CARL J., and IVAN A. SAG. 1994. *Head-driven phrase structure grammar*. (Studies in contemporary linguistics.) Chicago: University of Chicago Press.

PULLUM, GEOFFREY K. 2003. Learnability: Mathematical aspects. *Oxford international encyclopedia of linguistics*, 2nd edn., ed. by William J. Frawley, 431–34. Oxford: Oxford University Press.

PULLUM, GEOFFREY K. 2007. The evolution of model-theoretic frameworks in linguistics. In Rogers & Kepser, 1–10. Online: http://cs.earlham.edu/esslli07mts/mtspapers/pullum.pdf.

PULLUM, GEOFFREY K. 2013. The central question in comparative syntactic metatheory. *Mind & Language* 28.4.392–420.

PULLUM, GEOFFREY K., and BARBARA C. SCHOLZ. 2001. On the distinction between generative-enumerative and model-theoretic syntactic frameworks. *Logical Aspects of Computational Linguistics: 4th international conference* (Lecture notes in computer science 2099), ed. by Philippe de Groote, Glyn Morrill, and Christian Retor, 17–43. Berlin: Springer.

RAUH, GISA. 2013. Linguistic categories and the syntax-semantics interface: Evaluating competing approaches. Wuppertal: Universität Wuppertal, MS.

REIS, MARGA. 1974. Syntaktische Hauptsatzprivilegien und das Problem der deutschen Wortstellung. *Zeitschrift für Germanistische Linguistik* 2.3.299–327.

RICHTER, FRANK. 2007. Closer to the truth: A new model theory for HPSG. In Rogers & Kepser, 101–10. Online: http://cs.earlham.edu/esslli07mts/mtspapers/richter.pdf.

ROGERS, JAMES, and STEPHAN KEPSER (eds.) 2007. *Model-Theoretic Syntax at 10: Proceedings of the ESSLLI 2007 MTS@10 Workshop*, August 13–17, Trinity College Dublin. Online: http://cs.earlham.edu/esslli07mts/.

SAG, IVAN A. 1997. English relative clause constructions. *Journal of Linguistics* 33.2.431–84.

SAG, IVAN A. 2010. English filler-gap constructions. *Language* 86.3.486–545.

SAG, IVAN A., and THOMAS WASOW. 2011. Performance-compatible competence grammar. *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, ed. by Robert Borsley and Kersti Börjars, 359–77. Oxford: Wiley-Blackwell.

SAG, IVAN A.; THOMAS WASOW; and EMILY M. BENDER. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI lecture notes 152.) Stanford, CA: CSLI Publications.

SAUERLAND, ULI, and PAUL ELBOURNE. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33.2.283–319.

SCHÄFER, ROLAND, and FELIX BILDHAUER. 2012. Building large corpora from the web using a new efficient tool chain. *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, 486–93.

STABLER, EDWARD P. 2001. Minimalist grammars and recognition. *Linguistic form and its computation* (Studies in computational linguistics 1), ed. by Christian Rohrer, Antje Rossdeutscher, and Hans Kamp, 327–52. Stanford, CA: CSLI Publications.

STABLER, EDWARD P. 2010. After government and binding theory. *Handbook of logic and language*, 2nd edn., ed. by Johan van Benthem and Alice ter Meulen, 395–414. Amsterdam: Elsevier.

STABLER, EDWARD P. 2011. Computational perspectives on minimalism. *The Oxford handbook of linguistic minimalism*, ed. by Cedric Boeckx, 616–41. Oxford: Oxford University Press.

STEEDMAN, MARK J. 1991. Structure and intonation. *Language* 67.2.260–96.

STEEDMAN, MARK J. 2000. *The syntactic process*. (Language, speech, and communication.) Cambridge, MA: MIT Press.

STEEDMAN, MARK J., and JASON BALDRIDGE. 2006. Combinatory categorial grammar. *Encyclopedia of language and linguistics*, 2nd edn., ed. by Keith Brown, 610–21. Oxford: Elsevier.

STERNEFELD, WOLFGANG, and FRANK RICHTER. 2012. Wo stehen wir in der Grammatiktheorie?— Bemerkungen anläßlich eines Buchs von Stefan Müller. *Zeitschrift für Sprachwissenschaft* 31.2.263–91.

THIERSCH, CRAIG L. 1978. *Topics in German syntax*. Cambridge, MA: MIT dissertation.

TIMBERLAKE, ALAN. 1982. The impersonal passive in Lithuanian. *Berkeley Linguistics Society* 8.508–24.

TOMASELLO, MICHAEL. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

VAN NOORD, GERTJAN, and GOSSE BOUMA. 1994. The scope of adjuncts and the processing of lexical rules. *Proceedings of the International Conference on Computational Linguistics (COLING '94)* 15.250–56. Online: http://grid.let.rug.nl/~vannoord/papers/coling94.ps.gz.

VEENSTRA, METTINA JOLANDA ARNOLDINA. 1998. *Formalizing the minimalist program*. Groningen: University of Groningen dissertation.

VOGEL, RALF. 2001. Case conflict in German free relative constructions: An optimality theoretic treatment. *Competition in syntax*, ed. by Gereon Müller and Wolfgang Sternefeld, 341–75. Berlin: Mouton de Gruyter.

WECHSLER, STEPHEN MARK. 2008. A diachronic account of English deverbal nominals. *West Coast Conference on Formal Linguistics (WCCFL)* 26.498–506.

WEXLER, KENNETH, and PETER W. CULICOVER. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

WILLIAMS, EDWIN. 1994. Remarks on lexical knowledge. *Lingua* 92.7–34.

Arbeitsgruppe Deutsche Grammatik
Institut für Deutsche und Niederländische Philologie
Freie Universität Berlin
Habelschwerdter Allee 45
14 195 Berlin, Germany
[Stefan.Mueller@fu-berlin.de]